

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 14, 2004, 15:03:35 ; Search time 13362 Seconds  
(without alignments)  
6967.582 Million cell updates/sec

Title: US-09-813-329-5  
Perfect score: 2148  
Sequence: 1 ggcacggcgacgacgt.....aaaaaaaaaactcgag 2148

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 100%  
Listing first 45 summaries

Database :

GenEmbl.\*

1: gb.ba.\*

2: gb.htg.\*

3: gb.in.\*

4: gb.om.\*

5: gb.ov.\*

6: gb.pat.\*

7: gb.ph.\*

8: gb.pl.\*

9: gb.pr.\*

10: gb.ro.\*

11: gb.scs.\*

12: gb.ey.\*

13: gb.un.\*

14: gb.vi.\*

15: em.ba.\*

16: em.fun.\*

17: em.hum.\*

18: em.in.\*

19: em.mu.\*

20: em.om.\*

21: em.or.\*

22: em.ov.\*

23: em.pat.\*

24: em.ph.\*

25: em.pl.\*

26: em.ro.\*

27: em.scs.\*

28: em.un.\*

29: em.vi.\*

30: em.htg.hum.\*

31: em.htg.inv.\*

32: em.htg.other.\*

33: em.htg.mus.\*

34: em.htg.pln.\*

35: em.htg.rod.\*

36: em.htg.mam.\*

37: em.htg.vrt.\*

38: em.ay.\*

39: em.htgo.hum.\*

40: em.htgo.mus.\*

41: em.htgo.other.\*

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query			ID	Description	
		Match	Length	DB			
1	2103.8	97.9	2165	3	BT001838	BT001838 Drosophil	
2	2101	97.8	2101	3	AF149799	AF149799 Drosophil	
3	2076.4	96.7	2159	3	AF521176	AF521176 Drosophil	
4	1656	77.1	1656	3	AB073865	AB073865 Drosophil	
5	1202	56.0	1248	3	AY115551	AY115551 Drosophil	
6	1149.4	53.5	1221	3	AY119233	AY119233 Drosophil	
7	798.8	37.2	55359	2	AC012854	AC012854 Drosophil	
C	798.8	37.2	158983	3	AC005974	AC005974 Drosophil	
	798.8	37.2	172904	3	AC007414	AC007414 Drosophil	
C	798.8	37.2	189620	3	AC090029	AC090029 Drosophil	
	798.8	37.2	275390	3	AE003831	AE003831 Drosophil	
11	798.8	37.2	275390	3	AE003831	AE003831 Drosophil	
12	67.2	3.1	2040	5	AB117518	AB117518 Theragra	
13	67.2	3.1	2235	5	AB117517	AB117517 Theragra	
14	66.8	3.1	390	6	AR135147	AR135147 Sequence	
15	66.4	3.1	2000	6	AX655393	AX655393 Sequence	
16	65.2	3.0	148198	2	BX276082	BX276082 Danio rerio	
C	64.4	3.0	215119	2	AC098082	AC098082 Rattus norvegicus	
	62.4	2.9	143586	5	AL627248	AL627248 Zebrafish	
C	18	62.4	2.9	148339	2	BX677670	BX677670 Danio rerio
C	20	60.8	2.8	232092	2	AC127800	AC127800 Rattus norvegicus
21	60.4	2.8	243597	2	AC099373	AC099373 Rattus norvegicus	
22	59.6	2.8	229583	10	AL844566	AL844566 Mouse DNA	
C	23	59.4	2.8	259427	2	AC137318	AC137318 Rattus norvegicus
C	24	59.4	2.8	284938	2	AC091335	AC091335 Rattus norvegicus
C	25	58.8	2.7	643	3	AY301008	AY301008 Pogonomys
26	58.6	2.7	234817	10	AL663048	AL663048 Mouse DNA	
27	58	2.7	3127	14	AF305694	AF305694 Kaposi's sarcoma	
28	56.8	2.6	206973	2	AC130594	AC130594 Rattus norvegicus	
29	56.8	2.6	247900	2	AC136408	AC136408 Rattus norvegicus	
C	30	56.2	2.6	223508	2	AR072952	AR072952 Sequence
	31	55.8	2.6	3489	6	AR072952	AR072952 Sequence
32	55.8	2.6	3489	6	AR254713	AR254713 Sequence	
33	55.8	2.6	3489	6	AR261360	AR261360 Sequence	
34	55.8	2.6	3489	6	AX107938	AX107938 Sequence	
35	55.8	2.6	3489	14	KSU52064	US2064 Kaposi's sarcoma	
C	36	55.8	2.6	32207	6	AR065852	AR065852 Sequence
C	37	55.8	2.6	32207	6	AR127850	AR127850 Sequence
C	38	55.8	2.6	32207	6	AR194752	AR194752 Sequence
C	39	55.8	2.6	32207	6	BD190717	BD190717 Unique assembly
C	40	55.8	2.6	137508	14	KSU75698	U75698 Kaposi's sarcoma
41	55.6	2.6	620	9	BC035567	BC035567 Homo sapiens	
42	55.6	2.6	1847	3	AY069729	AY069729 Drosophil	
43	55.6	2.6	2931	14	AF192756	AF192756 Kaposi's sarcoma	
44	55.4	2.6	1779	3	AB104617	AB104617 Epididymus	
C	45	55	2.6	88549	3	AC116979	AC116979 Dictyostelium

ALIGNMENTS

RESULT 1  
BT001838 2165 bp mRNA linear INV 15-NOV-2002  
LOCUS Drosophila melanogaster RH51659 full ineert cdna.  
DEFINITION BT001838  
ACCESSION  
VERSION BT001838.1 GI:25013017  
KEYWORDS FLI\_CDNA.  
SOURCE Drosophila melanogaster (fruit fly)  
ORGANISM Drosophila melanogaster  
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
Ephydroidea; Drosophilidae; Drosophila.  
REFERENCE 1 (bases 1 to 2165)  
AUTHORS Scapleton,M., Broketein,P., Hong,L., Agbayani,A., Carlson,J.,  
Champe,M., Chavez,C., Dorsett,V., Dresnek,D., Farfan,D., Frise,E.,

George, R., Gonzalez, M., Guarin, H., Krommiller, B., Li, P., Liao, G., Miranda, A., Mungall, C. J., Nunco, J., Pacleb, J., Paragas, V., Park, S., Patel, S., Phouanavong, S., Wan, K., Yu, C., Lewis, S. E., Rubin, G. M. and Celisner, S.

#### TITLE

#### JOURNAL

Submitted (15-NOV-2002) Berkeley Drosophila Genome Project, Lawrence Berkeley National Laboratory, One Cyclotron Road, Berkeley, CA 94720, USA

#### COMMENT

Sequence submitted by:

Berkeley Drosophila Genome Project

Lawrence Berkeley National Laboratory

Berkeley, CA 94720

This clone was sequenced as part of a high-throughput process to sequence clones from Drosophila Gene Collection 1 (Rubin et al., Science 2000). The sequence has been subjected to integrity checks for sequence accuracy, presence of a polyA tail and contiguity within 100 kb in the genome. Thus we believe the sequence to reflect accurately this particular cDNA clone. However, there are artifacts associated with the generation of cDNA clones that may have not been detected in our initial analyses such as internal priming, priming from contaminating genomic DNA, retained introns due to reverse transcription of unspliced precursor RNAs, and reverse transcriptase errors that result in single base changes. For further information about this sequence, including its location and relationship to other sequences, please visit our Web site (<http://fruitfly.berkeley.edu>) or send email to [cdna@fruitfly.berkeley.edu](mailto:cdna@fruitfly.berkeley.edu).

#### FEATURES

##### source

1..2165  
/organism="Drosophila melanogaster"  
/mol\_type="mRNA"  
/strain="Y; cn bw sp"  
/db\_xref="taxon:7227"

##### misc\_feature

1..2152  
/note="sim4 alignment with AE003831 (46B13-46E4)"

648..1877

##### CDS

/note="Longest ORF"  
/codon\_start=1  
/product="RH51659p"  
/protein\_id="AA0171595.1"  
/db\_xref="GI:25013018"

/translation="MTAEILKPFITPSANDGPPAKATATAQRTRTOLIPLVLGFI  
IGLVVAIIALITWQTRVRSKLDKSLKRVVNDLQRLGNYLDDEFQKEYEN  
ALIDPKKVDLTDEDDDDGDLSDADDDDDVSVDVGDVDEYDMLNKL  
NAHTGPTTSETAEGGETDSASNDNDVDFDFTSYNAHKKOKERSRSDVNRN  
EEONIQGNHTELOKSNEATSKERPAPLHRRMRSHRHLVRKARSDSRPAHF  
HLSSRRHQSMGTHGMYIGNERNISQGHFQTRDGLVTVTNIGLYIYVIAQICYN  
SHDQNGFIVFGDTPFLQCLNTVPTNMPHKVHTCHTSLHLERNERIHLKDIHNDNRN  
AVLREGNNRSYFGIFKV"

#### ORIGIN

Query Match 97.9%; Score 2103.8; DB 3; Length 2165;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 2105; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 5 CGAGCGGAAACGACGCTTTAAAGTGAGAAAAGAAACCGGTAAATCAGAGATCCCAAGCAAG 64  
Db 19 CGACGCGAAACGAGCTTTAAAGTGAGAAAAGAAACCGGTAAATCAGAGATCCCAAGCAAG 78

Qy 65 CGCGTGGTGATGATAGCGAAGAAAAAGCTATCCGTTTCAGTTTAACTACTTACCTACCAAG 124  
Db 79 CGCGTGGTGATGATAGCGAAGAAAAAGCTATCCGTTTCAGTTTAACTACTTACCTACCAAG 138

Qy 125 ATTGAATTTCCGCAATCGGCAAAATTTACTAAAAATACATAAGTCGAATCGCTCCACTGTGT 184  
Db 139 ATTGAATTTCCGCAATCGGCAAAATTTACTAAAAATACATAAGTCGAATCGCTCCACTGTGT 198

Qy 185 GTTGTTGTTTTTTTTTTTTTTTTTTTGGTTTCGCTGTGCTTTATCGCAAAACAAGAACTGAT 244  
Db 199 GTTGTTGTTTTTTTTTTTTTTTTTTTGGTTTCGCTGTGCTTTATCGCAAAACAAGAACTGAT 258

Qy 245 AAAACTAGAAAATATCTTGAGAAACTGTTTCGCGCTTTCTTTTGGCTTAATTTGCGGATC 304  
Db 1339 AGGAAAAGTGTCCAAATGAGGCAACTTCCAAAGAGAGACCTTGACCACTTCCACCGGTC 1398

Db 259 ABAACTAGAAAATATCTTGAGAAACTGTTGTTTTCGCGCTTTCTTTTGGCTAATTTGCCGATC 318  
Qy 305 GCGGAGAGAAAAACAAGCAGTGTAGACAAACAAGTGTGGTAAATACAAATCTGAAAAAGGCA 364  
Db 319 GCGGAGAGAAAAACAAGCAGTGTAGACAAACAAGTGTGGTAAATACAAATCTGAAAAAGGCA 378  
Qy 365 CCATCAGCAGCCCGAGGGTTTATCTATATAGATGTGCGAGCTTATCATCTCATCTGCTGTC 424  
Db 379 CCATCAGCAGCCCGAGGGTTTATCTATATAGATGTGCGAGCTTATCATCTCATCTGCTGTC 438  
Qy 425 TGTGAGGTTTCTTCTGTGCTGCTAGTAGTATCTTAAATACATAGAGTGTGTTTCATATAA 484  
Db 439 TGTGAGGTTTCTTCTGTGCTGCTAGTAGTATCTTAAATACATAGAGTGTGTTTCATATAA 498  
Qy 485 GTGCGACAAAGCTCGATTGGAACACGCTGTGAGTGCCTTTGAGTGGGTGGGCAAGATCG 544  
Db 499 GTGCGACAAAGCTCGATTGGAACACGCTGTGAGTGCCTTTGAGTGGGTGGGCAAGATCG 558  
Qy 545 TCATCATCATCATCTGCTGCTCATTCACACAGAAATCAGCATCAGCATCTGGAGCCCGGT 604  
Db 559 TCATCATCATCATCTGCTGCTCATTCACACAGAAATCAGCATCAGCATCTGGAGCCCGGT 618  
Qy 605 TGCTCTAAGATCCCGAGTGTTCATCAATATATGACTGCCGAGACCCCTCAAGCCGTTTATAA 664  
Db 619 TGCTCTAAGATCCCGAGTGTTCATCAATATATGACTGCCGAGACCCCTCAAGCCGTTTATAA 678  
Qy 665 CGCCACAGGTGCCACAGATGATGTTTTCGGGCCAAAGCGACACGCGGACCGGCC 724  
Db 679 CGCCACAGGTGCCACAGATGATGTTTTCGGGCCAAAGCGACACGCGGACCGGCC 738  
Qy 725 AGCGACGACCCCGCAGCTGATCCCTCGTTTGGGGTTCATCGGTCTGGGGCTGGTCG 784  
Db 739 AGCGACGACCCCGCAGCTGATCCCTCGTTTGGGGTTCATCGGTCTGGGGCTGGTCG 798  
Qy 785 TTGCCATTCTCGACCTAACGATCTGGCAGACAAACGCTGTATCGCATCTGGACAAGAGAC 844  
Db 799 TTGCCATTCTCGACCTAACGATCTGGCAGACAAACGCTGTATCGCATCTGGACAAGAGAC 858  
Qy 845 TGAAGAGCTGAGCGAGTCTCGATATCTCCAGCAGCGTTTGGGCATTAACCTATCTCG 904  
Db 859 TGAAGAGCTGAGCGAGTCTCGATATCTCCAGCAGCGTTTGGGCATTAACCTATCTCG 918  
Qy 905 ACAGTTTCACAGATTTCAAAAAGGAGTAGCAGAAATGCCCTCATCGACTATCAAAAAAGG 964  
Db 919 ACAGTTTCACAGATTTCAAAAAGGAGTAGCAGAAATGCCCTCATCGACTATCAAAAAAGG 978  
Qy 965 TGGATGGCTCTACGATAGAGGAGCAGCAGATGGCGATGGTCTGGAATCCATTGCGG 1024  
Db 979 TGGATGGCTCTACGATAGAGGAGCAGCAGATGGCGATGGTCTGGAATCCATTGCGG 1038  
Qy 1025 ACAGCAGGACGACGAGTGTAGCTATAGCTGTGATGATGTTTGGCGGAGACTAGGAG 1084  
Db 1039 ACAGCAGGACGACGAGTGTAGCTATAGCTGTGATGATGTTTGGCGGAGACTAGGAG 1098  
Qy 1085 ACTTACACCGATATGTTAAATAAATCAACATGCAATACCGGCACCGCCCATCTG 1144  
Db 1099 ACTTACACCGATATGTTAAATAAATCAACATGCAATACCGGCACCGCCCATCTG 1158  
Qy 1145 AGACCACTCTGAGGGCGAGGGCGAGCAGCAGTGTGATCTTCAGCTCAAAATGATGACA 1204  
Db 1159 AGACCACTCTGAGGGCGAGGGCGAGCAGCAGTGTGATCTTCAGCTCAAAATGATGACA 1218  
Qy 1205 ATGTGTTGATGACTTTTACAGCTCAATGCCCCACAAAAGAGAGCAGGAGAAATCTC 1264  
Db 1219 ATGTGTTGATGACTTTTACAGCTCAATGCCCCACAAAAGAGAGCAGGAGAAATCTC 1278  
Qy 1265 GCTCGATTGCCGATGTACCAATGAGGAGCAATATTTCAAGGAAATCACACAGAGCTTC 1324  
Db 1279 GCTCGATTGCCGATGTACCAATGAGGAGCAATATTTCAAGGAAATCACACAGAGCTTC 1338  
Qy 1325 AGGAAAAGTGTCCAAATGAGGCAACTTCCAAAGAGAGCGCTGACCACTTACCAACCGTC 1384  
Db 1339 AGGAAAAGTGTCCAAATGAGGCAACTTCCAAAGAGAGCGCTGACCACTTACCAACCGTC 1398

QY	1385	GCAGAATGCATTTCCGCCATCGCCACCTCTAGTCCGCAAGCCAGATCCGAGACTCGA	1444		
DB	1399	GCAGAATGCATTTCCGCCATCGCCACCTCTAGTCCGCAAGCCAGATCCGAGACTCGA	1458		
QY	1445	GGCAGCAGGCCATTTCCACTTGAGCAGCAGGCGCGCTCACCAAGGAAGTATGGCGTACC	1504		
DB	1459	GGCAGCAGGCCATTTCCACTTGAGCAGCAGGCGCGCTCACCAAGGAAGTATGGCGTACC	1518		
QY	1505	ATCGAGATATGTACATAGGAAATGATAACGAGAGAAACTCTTATCAGGGAACCTTTCAAA	1564		
DB	1519	ATCGAGATATGTACATAGGAAATGATAACGAGAGAAACTCTTATCAGGGAACCTTTCAAA	1578		
QY	1565	CGCGCATGGCGTCTTGACGGTGACCAATACAGGCCTATATTACGTATACGCCAGATAT	1624		
DB	1579	CGCGCATGGCGTCTTGACGGTGACCAATACAGGCCTATATTACGTATACGCCAGATAT	1638		
QY	1625	GCTACAACAACTCGCAGACAGAAACGGATTTATTCGCTTTTCAAGGAGACACTCCATTCC	1684		
DB	1639	GCTACAACAACTCGCAGACAGAAACGGATTTATTCGCTTTTCAAGGAGACACTCCATTCC	1698		
QY	1685	TGCAGTGTCTTGAACACGGTGTCCCAACCAACATAGCCACATAAGGTGCACACCTGCCACAGA	1744		
DB	1699	TGCAGTGTCTTGAACACGGTGTCCCAACCAACATAGCCACATAAGGTGCACACCTGCCACAGA	1758		
QY	1745	GTGGTCTGATCCACTTGAACGAAACGAGAGGATCCATCTGAAGGACATTCACAACGATC	1804		
DB	1759	GTGGTCTGATCCACTTGAACGAAACGAGAGGATCCATCTGAAGGACATTCACAACGATC	1818		
QY	1805	GCAATGCAAGTCTGCGGAGGGAAACCAACGAAAGTACTTTGGCATCTTCAAGGTGTAAA	1864		
DB	1819	GCAATGCAAGTCTGCGGAGGGAAACCAACGAAAGTACTTTGGCATCTTCAAGGTGTAAA	1878		
QY	1865	TTGGAGAGATTATCCCGGTGAGAGATGAAATACCAAGTTTAAAGCTTTTGTCCCGCGCAC	1924		
DB	1879	TTGGAGAGATTATCCCGGTGAGAGATGAAATACCAAGTTTAAAGCTTTTGTCCCGCGCAC	1938		
QY	1925	TGCTCGTGAATGCGATTATCGCCAGCGTGAATCCATTAGTTTGGTAGTACTAGTCTTTAG	1984		
DB	1939	TGCTCGTGAATGCGATTATCGCCAGCGTGAATCCATTAGTTTGGTAGTACTAGTCTTTAG	1998		
QY	1985	TCACTCCAAACCTAATCTCAATCGGAATCGTGCACTATGCAATAGTCAGAAGACGGAGGA	2044		
DB	1999	TCACTCCAAACCTAATCTCAATCGGAATCGTGCACTATGCAATAGTCAGAAGACGGAGGA	2058		
QY	2045	AAATCATATTTATTTTGTATATCTCGTTCGACTCTAAAAGGTGAATAAAAAATATATGTA	2104		
DB	2059	AAATCATATTTATTTTGTATATCTCGTTCGACTCTAAAAGGTGAATAAAAAATATATGTA	2118		
QY	2105	GCTATTA 2111			
DB	2119	GCTATTA 2125			
RESULT 2	AF149799	2101 bp	mRNA	linear	INV 02-JAN-2003
LOCUS	Drosophila melanogaster	tumor necrosis factor family member DT1			
DEFINITION	(dt1) mRNA, complete cds.				
ACCESSION	AF149799				
VERSION	AF149799.1	GI:27462085			
KEYWORDS					
SOURCE	Drosophila melanogaster	(fruit fly)			
ORGANISM	Drosophila melanogaster				
	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;				
	Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;				
	Ephydroidea; Drosophilidae; Drosophila.				
	1 (bases 1 to 2101)				
REFERENCE	Inohara,N. and Nunez,G.				
AUTHORS	DT1, a Drosophila tumor necrosis factor family member				
TITLE	Unpublished				
JOURNAL					
REFERENCE	2 (bases 1 to 2101)				
AUTHORS	Inohara,N. and Nunez,G.				

TITLE	Direct Submission
JOURNAL	Submitted (10-MAY-1999) Department of Pathology, Comprehensive Cancer Center, 1500 E. Medical Center Dr., Ann Arbor, MI 48109, USA
FEATURES	Location/Qualifiers
source	1. .2101
	/organism="Drosophila melanogaster"
	/mol_type="mRNA"
	/db_xref="taxon:7227"
gene	1. .2101
	/gene="dt1"
CDS	625..1854
	/gene="dt1"
	/codon_start=1
	/product="tumor necrosis factor family member DT1"
	/protein_id="AA015310.1"
	/db_xref="GI:27462086"
	/translation="MTAETLKPFTPTSANDGFFAKATSTATAQRTRQLPLVLGFG IGLVVAILATIWQTRVSHLDKELSKRVVDNIQRLGINYLDEPDEFQKEYEN ALIDYPKVDGLTDEDDDDGGLSDIADDEDDDVSYSSVDDVYGADYEDYDMLKLN NAHTGTPPTSETTAEGEGETDSASSISNDNDVDDFTSYNAHKQKRSIADVRN EEQNIQGNHTLEQKSSNEATSKESAPLHRRMRHSRHLVVRKARSERSRAAHP HLSSRRHQSGMGYHGDMYIGNDNERNYSQGHFOTRDGLVLTNTGLVYVYAOICYNH SHQNGEIVFGDTPPELQCLNTVPTNMPHKVHTCHTSLGHLERNERIHLKDIHNRN AVLREGNNRSYFGIFKV"
	2080..2085
polyA_signal	/gene="dt1"
	2101
polyA_site	/gene="dt1"
ORIGIN	
	Query Match 97.8%; Score 2101; DB 3; Length 2101; Best Local Similarity 100.0%; Pred. No. 0; Matches 2101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy	10 CGAAGCGACGTTTAAAGTGAGAAAAAGAAACGGGTAATACAGAGATCCCAAGCAAGCGGT 69
Db	1 CGAAGCGACGTTTAAAGTGAGAAAAAGAAACGGGTAATACAGAGATCCCAAGCAAGCGGT 60
Qy	70 GCGTGCATGATAGCGAAGAAAAAGCTATCCGTTTCAGTTAACTACTTACCAAGATTGA 129
Db	61 GCGTGCATGATAGCGAAGAAAAAGCTATCCGTTTCAGTTAACTACTTACCAAGATTGA 120
Qy	130 ATTTGCGCATCGGGCAAAATTAATAAATAACATAAGTGCACCTCGTCCACTGTGTGTGT 189
Db	121 ATTTGCGCATCGGGCAAAATTAATAAATAACATAAGTGCACCTCGTCCACTGTGTGTGT 180
Qy	190 GTTTTTTTTTTTTTTTTTTGGTTTTTGGCTGTGCGCTTTATCGCAAAACAGAACTGATAAAC 249
Db	181 GTTTTTTTTTTTTTTTTTTGGTTTTTGGCTGTGCGCTTTATCGCAAAACAGAACTGATAAAC 240
Qy	250 TAGAAAAATATCTTGAGAAACTTGTTCGCGCTTTCTTTTGTAAATTCGCGATCGCGGA 309
Db	241 TAGAAAAATATCTTGAGAAACTTGTTCGCGCTTTCTTTTGTAAATTCGCGATCGCGGA 300
Qy	310 AGAGAAAAACAAGCAGTAGACAAAAACAAGTGTGGTAAATACAACTTGAAGGGCACCATC 369
Db	301 AGAGAAAAACAAGCAGTAGACAAAAACAAGTGTGGTAAATACAACTTGAAGGGCACCATC 360
Qy	370 AGCAGCCCGAGGGTTTATCTATATAGATGTCGAGCTTATCATCTCATGCTGTCTGTGA 429
Db	361 AGCAGCCCGAGGGTTTATCTATATAGATGTCGAGCTTATCATCTCATGCTGTCTGTGA 420
Qy	430 GGTTGTTCTGTGCTCGTGTAGTATCTTAAATACATAGAGTGTGTTTCATATAAAGTGGC 489
Db	421 GGTTGTTCTGTGCTCGTGTAGTATCTTAAATACATAGAGTGTGTTTCATATAAAGTGGC 480
Qy	490 ACAAGAGCTCGATTGGAAACAGCTGTTCGAGTGCCTTCGAGTGGGTGGGCAAGATCGTCATC 549
Db	481 ACAAGAGCTCGATTGGAAACAGCTGTTCGAGTGCCTTCGAGTGGGTGGGCAAGATCGTCATC 540
Qy	550 ATCATCATCGTGTCTATTATCAACAGAAATCAGCATCAGCATCTGGAGGCCCGGTGTGCTC 609
Db	541 ATCATCATCGTGTCTATTATCAACAGAAATCAGCATCAGCATCTGGAGGCCCGGTGTGCTC 600

```
Qy 610 TAAGATCCCAGTGTTCATCAATATGACTGCCGAGACCCCTCAAGCCGTTTATTAACGCCA 669
Db 601 TAAGATCCCAGTGTTCATCAATATGACTGCCGAGACCCCTCAAGCCGTTTATTAACGCCA 660
Qy 670 ACGAGTGCACACGATGATGGTTCCTCCGCCAAAGCGACGACGACGCGCCGACCGCCAGCGA 729
Db 661 ACGAGTGCACACGATGATGGTTCCTCCGCCAAAGCGACGACGACGCGCCGACCGCCAGCGA 720
Qy 730 CGCACCCGCGAGCTGATCCCCCTGGTTTGGGTTTCATCGGTCTGGGCTGGTCTGGTGGC 789
Db 721 CGCACCCGCGAGCTGATCCCCCTGGTTTGGGTTTCATCGGTCTGGGCTGGTCTGGTGGC 780
Qy 790 ATTCTCGCATTAACGATCTGCAGACACGCGTGTATCGCATCTGGACAGGACCTGAAG 849
Db 781 ATTCTCGCATTAACGATCTGCAGACACGCGTGTATCGCATCTGGACAGGACCTGAAG 840
Qy 850 AGCCTGAAGCAGTGGTTCGATAATCTCCAGCAGAGGTTTGGGCATAAACTATCTGGACGAG 909
Db 841 AGCCTGAAGCAGTGGTTCGATAATCTCCAGCAGGTTTGGGCATAAACTATCTGGACGAG 900
Qy 910 TTCGACGAGTTCGAAAGGAGTACGAGAAATGCCCTCATCGCATATCAAAAAAGGTTGGAT 969
Db 901 TTCGACGAGTTCGAAAGGAGTACGAGAAATGCCCTCATCGCATATCAAAAAAGGTTGGAT 960
Qy 970 GGCCTCAGCGATGAGGAGGACGACGATGGCGATGGTCTGGATTCCATTTGGCGGACGAC 1029
Db 961 GGCCTCAGCGATGAGGAGGACGACGATGGCGATGGTCTGGATTCCATTTGGCGGACGAC 1020
Qy 1030 GAGCAGCAGCGTTAGCTATAGTCTGTGGATGATGTGGCGCAGACTACGAGGACTAC 1089
Db 1021 GAGCAGCAGCGTTAGCTATAGTCTGTGGATGATGTGGCGCAGACTACGAGGACTAC 1080
Qy 1090 ACCGATATGTTAAATRAACTCAACATGCACATACCGGCACGACCGCCACATCTGAGACC 1149
Db 1081 ACCGATATGTTAAATRAACTCAACATGCACATACCGGCACGACCGCCACATCTGAGACC 1140
Qy 1150 ACTGCTGAGGCGGAGGCGAGACGACGATGTCATCTCAGCTCAATATGATGACAAATGTG 1209
Db 1141 ACTGCTGAGGCGGAGGCGAGACGACGATGTCATCTCAGCTCAATATGATGACAAATGTG 1200
Qy 1210 TTCGATGACTTTACGACTACAATGCCACAAAAGAGCAGGAGAGAAAATCTCGCTCG 1269
Db 1201 TTCGATGACTTTACGACTACAATGCCACAAAAGAGCAGGAGAGAAAATCTCGCTCG 1260
Qy 1270 ATTGCCGATGTACCAATGAGGACGAGAAATTCAGGAAATTCACAGAGACTTCAGGAA 1329
Db 1261 ATTGCCGATGTACGCAATGAGGACGAGAAATTCAGGAAATTCACAGAGACTTCAGGAA 1320
Qy 1330 AAGTCATCCATGAGGCAACTTCCAAAGAGAGCCCTGCACCCTTCACCACCGTCGCAGA 1389
Db 1321 AAGTCATCCATGAGGCAACTTCCAAAGAGAGCCCTGCACCCTTCACCACCGTCGCAGA 1380
Qy 1390 ATGCATTTCCGCCCATCGCCACCTCTAGTCGCGAAAGCCAGATCCGAGGACTCGAGGCCA 1449
Db 1381 ATGCATTTCCGCCCATCGCCACCTCTAGTCGCGAAAGCCAGATCCGAGGACTCGAGGCCA 1440
Qy 1450 GCAGCCCATTTCCATTGAGCAGCAGCGGGGCTCACCAAGGAAGTATGGCTACCATGGA 1509
Db 1441 GCAGCCCATTTCCATTGAGCAGCAGCGGGGCTCACCAAGGAAGTATGGCTACCATGGA 1500
Qy 1510 GATATGTACATAGGAATGATTAACGAGAGAACTCTTATCAGGACACATTTTCAAAACGCGC 1569
Db 1501 GATATGTACATAGGAATGATTAACGAGAGAACTCTTATCAGGACACATTTTCAAAACGCGC 1560
Qy 1570 GATGGCTCTTGACGTTGACCAATACAGGCTATATTAACGTATACGCCCCAGATATGCTAC 1629
Db 1561 GATGGCTCTTGACGTTGACCAATACAGGCTATATTAACGTATACGCCCCAGATATGCTAC 1620
Qy 1630 AACAACTCGCAGCAGCAGAGAGGATTTATCGTCTTTCAAGGACACATCTCATTTCTGCGAG 1689
Db 1621 AACAACTCGCAGCAGCAGAGAGGATTTATCGTCTTTCAAGGACACATCTCATTTCTGCGAG 1680
```

```
Qy 1690 TGCTTGAACACGGTGCCCAACCAATGCCACATTAAGGTGCACACCTTGCACACAGTGGT 1749
Db 1681 TGCTTGAACACGGTGCCCAACCAATGCCACATTAAGGTGCACACCTTGCACACAGTGGT 1740
Qy 1750 CTGATCCACCTGGAACGAAACGAGAGGATCCATCTGAAGGACATTCACACGATCGCAAT 1809
Db 1741 CTGATCCACCTGGAACGAAACGAGAGGATCCATCTGAAGGACATTCACACGATCGCAAT 1800
Qy 1810 GCAGTTCTGCGGAGGAGAAACAAACCGAAGCTACTTTGGCATCTTCAAGGTGTAATTTGGA 1869
Db 1801 GCAGTTCTGCGGAGGAGAAACAAACCGAAGCTACTTTGGCATCTTCAAGGTGTAATTTGGA 1860
Qy 1870 GAGATATCCCGGTGAGAGATGGAATACCAAGTTTAAGCTTTTGTCCCGCAGCTGCTC 1929
Db 1861 GAGATATCCCGGTGAGAGATGGAATACCAAGTTTAAGCTTTTGTCCCGCAGCTGCTC 1920
Qy 1930 GTGAATGCCGATTCATCGCCAGCGTGAATCCATTTAGTTTCGTAGTACCTTAGTCTTAGTCACT 1989
Db 1921 GTGAATGCCGATTCATCGCCAGCGTGAATCCATTTAGTTTCGTAGTACCTTAGTCTTAGTCACT 1980
Qy 1990 CCAAACTTAATCTCAATCGGAATCGTGCACTAGTTCAGAAAGCGGAGGAAAAATC 2049
Db 1981 CCAAACTTAATCTCAATCGGAATCGTGCACTAGTTCAGAAAGCGGAGGAAAAATC 2040
Qy 2050 ATATTTATTTTGTATATCTCGTTCCAGCTCTAAAGTGAATAAATAATATATAGTACTAT 2109
Db 2041 ATATTTATTTTGTATATCTCGTTCCAGCTCTAAAGTGAATAAATAATATATAGTACTAT 2100
Qy 2110 T 2110
Db 2101 T 2101

RESULT 3
AF521176
LOCUS
DEFINITION
AF521176
ACCESSION
AF521176.1
VERSION
AF521176.1
KEYWORDS
SOURCE
ORGANISM
Drosophila melanogaster (fruit fly)
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE
1 (bases 1 to 2159)
AUTHORS
Moreno,E., Yan,M. and Basler,K.
Evolution of TNF Signaling Mechanisms. JNK-Dependent Apoptosis
Triggered by Eiger, the Drosophila Homolog of the TNF Superfamily
Curt. Biol. 12 (14), 1263-1268 (2002)
JOURNAL
12176339
PUBMED
2 (bases 1 to 2159)
REFERENCE
Moreno,E., Yan,M. and Basler,K.
Direct Submission
Submitted (13-JUN-2002) Institut Molecular Biology, Uni Zurich,
Winterthurerstrasse 190, Zurich, Z 8057, Switzerland
JOURNAL
Location/Qualifiers
FEATURES
source
1..2159
/organism="Drosophila melanogaster"
/mol_type="mRNA"
/db_xref="taxon:7227"
641..1888
/codon_start=1
/product="eiger"
/protein_id="AAM76710.1"
/db_xref="GI:21717646"
translation="MTASTLKPFIPTTSANDDGFPAKATSTATAQRRRLPLVLGF
IGLGLVAILALTITWQTRVSHLDKSLKRVNLOQLGINVLDSEDFQKEVEN
ALIDYPKVDGTLDEEDDDGDLSDIADDDVDYSVSDVDVGADYEDYDMLKLN
NAHTGTTPTSETTAEGGETDSASNDNVDFDFTSINAHKKQKERSRSIADVRN
EQNTQGNHTLEQKESNEATSKESPAPLHRRHRMHSRHLRLKQKERSLSARSED
PPAAPHLSRRRHOGSGYHGVIGNDNERNYSYQGFOTRDTGLTNTVNTGLYVYA
QTCYNSHDQNGFIVFQGDPTPFLQCLNTVPTMPHKVHTCHTSGLIHLERNRIHLKD
```



ORIGIN	IHNDRNAVLRBEGNRSYFPGIFKV"	Query Match Best Local Similarity 96.7%; Score 2076.4; DB 3; Length 2159; Matches 2121; Conservative 0; Mismatches 6; Indels 21; Gaps 2;
Qy	5	CGAGGCGAAGCGGACGTTTAAAGTGAGAAAGAAACCGGTAAATCAGAGATCCCAGCAAG 64
Db	15	CGAAGCGAAACGGACGTTTAAAGTGAGAAAGAAACCGGTAAATCAGAGATCCCAGCAAG 74
Qy	65	CGCGTCGGTGCGATAGAGGAGAAAGAAAGCTATCCGTTTTCAGTTAACTACTTACCAAG 124
Db	75	CGCGTCGGTGCGATAGAGGAGAAAGAAAGCTATCCGTTTTCAGTTAACTACTTACCAAG 134
Qy	125	ATTGAAATTTCCGCATCGGCAAAATTAACATAAAGTGAACCTCGTCCACTGTGT 184
Db	135	ATCGAATTTCCGCATCGGCAAAATTAACATAAAGTGAACCTCGTCCACTGTGT 194
Qy	185	GTTGTGTTTTTTTTTTTTTTTTTTTTTTCGTCGTGCTTTATCGCAAAACGAACCTGAT 244
Db	195	GTTGTG---TTTTTTTTTTTTTTTTTTTTTTCGTCGTGCTTTATCGCAAAACGAACCTGAT 251
Qy	245	AAAACCTAGAAAATATCTTCAGAAAACCTGTTTTTCGCGCTTTCTTTTGCCTAATTCGCGATC 304
Db	252	AAAACCTAGAAAATATCTTCAGAAAACCTGTTTTTCGCGCTTTCTTTTGCCTAATTCGCGATC 311
Qy	305	CGGGAAGAGAAAAACAAGCAGTAGACAAAAACAAGTGTGTAAATACAATCTGAAAAGGGCA 364
Db	312	CGGGAAGAGAAAAACAAGCAGTAGACAAAAACAAGTGTGTAAATACAATCTGAAAAGGGCA 371
Qy	365	CCATCAGCAGCCGAGGGGTTTTTCTATATAGATCGCAGCTTATCATCTCATCTGCTGTC 424
Db	372	CCATCAGCAGCCGAGGGGTTTTTCTATATAGATCGCAGCTTATCATCTCATCTGCTGTC 431
Qy	425	TGTGAGGTTGTTCTGTGCTCGTAGTATCTTAAATACATAGAGTGTTCATATAAA 484
Db	432	TGTGAGGTTGTTCTGTGCTCGTAGTATCTTAAATACATAGAGTGTTCATATAAA 491
Qy	485	GTGCGCAAAAGCTCGAATCGAAACAGCTGTGAGTGCCCTTTGAGTGGGTGGCAAGATCG 544
Db	492	GTGCGCAAAAGCTCGAATCGAAACAGCTGTGAGTGCCCTTTGAGTGGGTGGCAAGATCG 551
Qy	545	TCATCATCATCATCGTCGTCATTATCAACAGAAATCAGCATCAGCATCTGGAGGCCCGGGT 604
Db	552	TCATCATCATCATCGTCGTCATTATCAACAGAAATCAGCATCAGCATCTGGAGGCCCGGGT 611
Qy	605	TGCTCTAAGATCCCGAGTGTTCATCAATATGACTGCGGAGACCCCTCAAGCCGTTTATAA 664
Db	612	TGCTCTAAGATCCCGAGTGTTCATCAATATGACTGCGGAGACCCCTCAAGCCGTTTATAA 671
Qy	665	CGCCAACGAGTGCCAACGATGATGTTTTTCGGCCAAAGCGACGACGCGACCGCC 724
Db	672	CGCCAACGAGTGCCAACGATGATGTTTTTCGGCCAAAGCGACGACGCGACCGCC 731
Qy	725	AGCGACGCCCGCCAGCTGATCCCGCTGTTTTGGGGTTTCATCGGTCTGGGGCTGGTCG 784
Db	732	AGCGACGCCCGCCAGCTGATCCCGCTGTTTTGGGGTTTCATCGGTCTGGGGCTGGTCG 791
Qy	785	TTGCCATTCTCGCACTTAAACGATCTGGCAGACAAACGCGTGTATCGCATCTGCGAAGGAGC 844
Db	792	TTGCCATTCTCGCACTTAAACGATCTGGCAGACAAACGCGTGTATCGCATCTGCGAAGGAGC 851
Qy	845	TGAAGAGCTGGAAGCGAGTCGTGCATATCTCCAGCAGCGTTTGGGCATAAATCTCTGG 904
Db	852	TGAAGAGCTGGAAGCGAGTCGTGCATATCTCCAGCAGCGTTTGGGCATAAATCTCTGG 911
Qy	905	ACGAGTTTCAACGAGTTTCCAAAAGGAGTACGAGAAATGCCCTCATCGCATCTCCAAAAGG 964
Db	912	ACGAGTTTCAACGAGTTTCCAAAAGGAGTACGAGAAATGCCCTCATCGCATCTCCAAAAGG 971
Qy	965	TGGATCGCCTCAACGATAGGAGGACGACGACGATGGCGATGGTCTGGATTTCCATTCGGG 1024



```
Db 1261 AGCAGCGCGCGTCAACAAGAAAGTATGGCTACCATGGAGATATGTACATAGGAAAT 1320
Qy 1528 GATACGAGAGAAACTCTTATCAGGAGACATTTCAAACGCGGATGGCGTCTTGACGGTG 1587
Db 1321 GATAACGAGAGAAACTCTTATCAGGAGACATTTCAAACGCGGATGGCGTCTTGACGGTG 1380
Qy 1588 ACCAATACAGGCTATATTACGTATACGCCAGATATGCTACAACAACCTCGCACGACCAG 1647
Db 1381 ACCAATACAGGCTATATTACGTATACGCCAGATATGCTACAACAACCTCGCACGACCAG 1440
Qy 1648 AACGATTATAGTCTTTCAAGGAGACACTCCATTCTCGAGTGTGAAACACGGTGGCC 1707
Db 1441 AACGATTATAGTCTTTCAAGGAGACACTCCATTCTCGAGTGTGAAACACGGTGGCC 1500
Qy 1708 ACCACATGCCACATAAGGTGCACACCTGCCACAGAGTGGTCTGATCCACTGGAAGA 1767
Db 1501 ACCACATGCCACATAAGGTGCACACCTGCCACAGAGTGGTCTGATCCACTGGAAGA 1560
Qy 1768 AACGAGAGGATCCATCTGAAGGACATTCACACGATCGCAATGCGAGTCTTGGCGGAGGA 1827
Db 1561 AACGAGAGGATCCATCTGAAGGACATTCACACGATCGCAATGCGAGTCTTGGCGGAGGA 1620
Qy 1828 AACACCGAAGCTACTTTGGCATCTTCAAGGTGTAA 1863
Db 1621 AACACCGAAGCTACTTTGGCATCTTCAAGGTGTAA 1656

RESULT 5
AY115551
LOCUS Drosophila melanogaster DARTH (darth) mRNA linear INV 01-AUG-2003
DEFINITION Drosophila melanogaster DARTH (darth) mRNA, complete cds.
ACCESSION AY115551
VERSION AY115551.1 GI:31321973
KEYWORDS
SOURCE
ORGANISM Drosophila melanogaster (fruit fly)
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 1248)
Kaupilla,S., Maaty,W.S., Chen,P., Tomar,R.S., Eby,M.T., Chapiro,J.,
Chew,S., Rathore,N., Zachariah,S., Sinha,S.K., Abrams,J.M. and
Chaudhary,P.M.
Eiger and its receptor, Wengen, comprise a TNF-like system in
Drosophila
Oncogene 22 (31), 4860-4867 (2003)
22775938
MEDLINE 12894227
PUBMED
REFERENCE 2 (bases 1 to 1248)
Chaudhary,P.M.
AUTHORS
TITLE Direct Submission
JOURNAL Submitted (28-MAY-2002) Internal Medicine, UT Southwestern Medical
Center, 5323 Harry Hines Blvd., Dallas, TX 75390, USA
Location/Qualifiers
FEATURES
1..1248
/organism="Drosophila melanogaster"
/mol_type="mRNA"
/db_xref="taxon:7227"
/chromosome="2"
/maps="2R"
1..1248
/gene="darth"
1..1248
/gene="darth"
/note="Drosophila aspartate rich TNF homolog; tumor
necrosis factor ligand family member"
/codon_start=1
/product="DARTH"
/protein_id="AA066763.1"
/db_xref="GI:31321974"
/translation="MTAETLKPFFITPTSANDDGFPAKATSTATQRRTRQLIPVLGF
IGLGLVVAIALTIQTTRVSHLDKELSLKRVVDNLQRLGINLYLDFDFQKEYEN
```

```
ALIDYPKVLDLDEDDDDGGLDLSIADDEDDVSYSSVDVGVADYEDYDTMLNKLK
NAHTGTPTSETTAEGEATSDASSASNDNDVDFDTSYNAHKKQKRSKSIADVRN
BEQNIQGNHTELOKSSNEATSKESPAFLHRRMRHSRHLHLVRKGESLKSARSEDS
RPAHFHLSRRRHQSGMGVHDWYIGNDAERNVSYQGHFOTRFDGLVTWNTGLYVYA
QICYNNSHDQNGFIVFOGDTFFLQCLNVTVPNMPHKVHTCHTSLGHLRNERIHLKD
IHNDRNAVLRGNRRNSYFGIPKV"
1..108
/misc_feature
/gene="darth"
/note="Region: intracellular domain"
misc_feature
109..177
/gene="darth"
/note="Region: transmembrane domain"
misc_feature
178..1245
/gene="darth"
/note="Region: extracellular domain"
misc_feature
433..438
/gene="darth"
/note="cleavage site"
ORIGIN
Query Match 56.0%; Score 1202; DB 3; Length 1248;
Best Local Similarity 98.6%; Pred. No. 1.6e-263;
Matches 1230; Conservative 0; Mismatches 0; Indels 18; Gaps 1;
Qy 634 ATGACTGCCGAGACCCCTCAAGCCGTTTATAACGCCAACGAGTGCACGATGATGTTT 693
Db 1 ATGACTGCCGAGACCCCTCAAGCCGTTTATAACGCCAACGAGTGCACGATGATGTTT 60
Qy 694 CCGGCCAAAGCGACGACGCGGACCGCCGAGGACGCGACCCGCGCAGCTGATCCCCCTG 753
Db 61 CCGGCCAAAGCGACGACGCGGACCGCCGAGGACGCGACCCGCGCAGCTGATCCCCCTG 120
Qy 754 GTTTTGGGGTTTCATCGGTCGCGGCTGGTCTGCTTCCATTCTCGCACTAACGATCTGGCAG 813
Db 121 GTTTTGGGGTTTCATCGGTCGCGGCTGGTCTGCTTCCATTCTCGCACTAACGATCTGGCAG 180
Qy 814 ACAACGGCTGTATCGCATCTGGACAAGGAGCTGAAGAGCCTGAAGCGAGTCTGTCGATAAT 873
Db 181 ACAACGGCTGTATCGCATCTGGACAAGGAGCTGAAGAGCCTGAAGCGAGTCTGTCGATAAT 240
Qy 874 CTCGACGACGCTTTGGGCGATAAACTATCTGACGAGTTCGACGAGTTCCTCAAAAGGAGTAC 933
Db 241 CTCGACGACGCTTTGGGCGATAAACTATCTGACGAGTTCGACGAGTTCCTCAAAAGGAGTAC 300
Qy 934 GAGAATGCCCTCATCGACTATCCAAAAGAGTGGATGGCCTCACGGATGAGGAGGACGAC 993
Db 301 GAGAATGCCCTCATCGACTATCCAAAAGAGTGGATGGCCTCACGGATGAGGAGGACGAC 360
Qy 994 GACGATGGCGATGGTCTGGATTCCATTGCGGACGACGAGGACGACGCTTAGCTATAGC 1053
Db 361 GACGATGGCGATGGTCTGGATTCCATTGCGGACGACGAGGACGACGCTTAGCTATAGC 420
Qy 1054 TCTGTGATGATGTTGGCGGACAGTACGAGACCTACACCGATATGTTAAATAACTCAAC 1113
Db 421 TCTGTGATGATGTTGGCGGACAGTACGAGACCTACACCGATATGTTAAATAACTCAAC 480
Qy 1114 AATCCATATCCGGCACCGCCACCATCTCAGACCACTGCTGAGGGGCGGGGAGAGACG 1173
Db 481 AATCCATATCCGGCACCGCCACCATCTCAGACCACTGCTGAGGGGCGGGGAGAGACG 540
Qy 1174 GACAGTGCATCTCAGCCCTCAAATCATGACAAATGTGTTTCGATGACTTTTACCAGCTACAAT 1233
Db 541 GACAGTGCATCTCAGCCCTCAAATCATGACAAATGTGTTTCGATGACTTTTACCAGCTACAAT 600
Qy 1234 GCCCAGAAAAAGAGAGGAGAGAAAAATCTCGCTCGATTGCCGATGTACGCAATGAGGAG 1293
Db 601 GCCCAGAAAAAGAGAGGAGAGAAAAATCTCGCTCGATTGCCGATGTACGCAATGAGGAG 660
Qy 1294 CAGAAATATTCAGGAAATCACACAGCTTCAGGAAAGTCAATCCATGAGGCACTTCC 1353
Db 661 CAGAAATATTCAGGAAATCACACAGCTTCAGGAAAGTCAATCCATGAGGCACTTCC 720
Qy 1354 AAGAGAGCCCTGCACCACTTTCACACCGTCGAGAAATGCATTTCCCGCCATCGCCACCTC 1413
```

Db 721 AAAGAGAGCGCTGCACCACTTCACCAAGTCGAGATGCATTCGCCCATGCCACCTC 780  
QY 1414 CTAGTCGGCAA-----AGCAGATCGAGAGCTCGAGGCGAGCAGCC 1455  
Db 781 CTAGTCGGCAAAGGTGAATCTCTTTTCAGCCAGATCCGAGGACTCGAGGCGAGCAGCC 840  
QY 1456 CATTTCCACTTGAGCAGCAGCGCGCTCACCAGGAAGTATGGCTTACCATGGAGATAG 1515  
Db 841 CATTTCCACTTGAGCAGCAGCGCGCTCACCAGGAAGTATGGCTTACCATGGAGATAG 900  
QY 1516 TACATAGGAATGATTAACAGAGAAACTCTTATCAGGACACTTTCAAACGCGCGATGCC 1575  
Db 901 TACATAGGAATGATTAACAGAGAAACTCTTATCAGGACACTTTCAAACGCGCGATGCC 960  
QY 1576 GTCTTCAGCGTGACCAATACAGCGCTTATTAAGTATACGCCAGATATGCTAACAAAC 1635  
Db 961 GTCTTCAGCGTGACCAATACAGCGCTTATTAAGTATACGCCAGATATGCTAACAAAC 1020  
QY 1636 TCGCAGCAGCAGAACGGATTTATCGTCTTTCAAGGAGACACTTCCATTCTCGAGTGTG 1695  
Db 1021 TCGCAGCAGCAGAACGGATTTATCGTCTTTCAAGGAGACACTTCCATTCTCGAGTGTG 1080  
QY 1696 AACACGGTCCCAACCAATGATGATGATGATGATGATGATGATGATGATGATGATG 1755  
Db 1081 AACACGGTCCCAACCAATGATGATGATGATGATGATGATGATGATGATGATGATG 1140  
QY 1756 CACTTGGAAACGAAACAGAGGATTCATCTGAGGACACTTCAACGATCGCAATGCGAGTT 1815  
Db 1141 CACTTGGAAACGAAACAGAGGATTCATCTGAGGACACTTCAACGATCGCAATGCGAGTT 1200  
QY 1816 CTGCGGAGGGAAACAAACCGAAGTACTTTGGCATCTTCAAGGTGTA 1863  
Db 1201 CTGCGGAGGGAAACAAACCGAAGTACTTTGGCATCTTCAAGGTGTA 1248

RESULT 6  
AY119233  
LOCUS Drosophila melanogaster SD18286 full insert cDNA. INV 16-JUN-2002  
DEFINITION  
ACCESSION AY119233  
VERSION AY119233.1 GI:21430829  
KEYWORDS FLI CDNA.  
SOURCE Drosophila melanogaster (fruit fly)  
ORGANISM Drosophila melanogaster  
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
Ephydroidea; Drosophilidae; Drosophila.  
1 (bases 1 to 1221)  
Stapleton, M., Brokstein, P., Hong, L., Agbayani, A., Carlson, J.,  
Champe, M., Chavez, C., Dorsett, V., Dresnek, D., Farfan, D., Frise, E.,  
George, R., Gonzalez, M., Guarin, H., Kronmiller, B., Li, P., Liao, G.,  
Miranda, A., Mungall, C.J., Nunco, J., Pacleb, J., Paragas, V., Park, S.,  
Patel, S., Phouanavong, S., Wan, K., Yu, C., Lewis, S.E., Rubin, G.M.,  
and Celisner, S.  
Direct Submission  
Submitted (07-JUN-2002) Berkeley Drosophila Genome Project,  
Lawrence Berkeley National Laboratory, One Cyclotron Road,  
Berkeley, CA 94720, USA  
Sequence submitted by:  
Berkeley Drosophila Genome Project  
Lawrence Berkeley National Laboratory  
Berkeley, CA 94720  
This clone was sequenced as part of a high-throughput process to  
sequence clones from Drosophila Gene Collection 1 (Rubin et al.,  
Science 2000). The sequence has been subjected to integrity checks  
for sequence accuracy, presence of a polyA tail and contiguity  
within 100 kb in the genome. Thus we believe the sequence to  
reflect accurately this particular cDNA clone. However, there are  
artifacts associated with the generation of cDNA clones that may  
have not been detected in our initial analyses such as internal  
priming, priming from contaminating genomic DNA, retained introns  
due to reverse transcription of unspliced precursor RNAs, and

reverse transcriptase errors that result in single base changes.  
For further information about this sequence, including its location  
and relationship to other sequences, please visit our web site  
(http://fruitfly.berkeley.edu) or send email to  
cdna@fruitfly.berkeley.edu.

FEATURES  
source  
1..1221  
/organism="Drosophila melanogaster"  
/mol\_type="mRNA"  
/db\_xref="taxon:7227"  
1..1221  
/gene="CG12919"  
/db\_xref="FLYBASE:FBgn0033483"  
170..955  
/notes="Longest ORF"  
/codon\_start=1  
/product="SD18286p"  
/protein\_id="AAM51093.1"  
/db\_xref="GI:21430830"  
/db\_xref="FLYBASE:FBgn0033483"  
/translation="MLNKLNNNAHTGTPITSETTAEGEGTDSASSASNDNVDDPFTS  
YNAKKQERKSRSTADVNERONIQNHTELQEKSSNEATSKSPAPLHRRMRHSR  
HRHLVVRKGSLSLARSERDPAAPHLSRRRHQSGMGTGMDTIGNDRNSYQGH  
FQTRDGLIVTNTGLYYVYQICYNNSHDQNGFI VFGDTPLQCLNTVPTNMPHKVH  
TCHTSLIHLERNRILHDKDIHNRNVLNREGRNSYFGIFKV"

ORIGIN

Query Match 53.5%; Score 1149.4; DB 3; Length 1221;  
Best Local Similarity 97.9%; Pred. No. 1.7e-251;  
Matches 1192; Conservative 0; Mismatches 6; Indels 19; Gaps 2;  
QY 931 TAGGAAATGCCCTCATCGACTATCAAAAAAGGTGGATGGCTCACGGATGAGGAGAC 990  
Db 6 TAGGAAATGCCCTCATCGACTATCC-CAAAAAGGTGGATGGCTCACGGATGAGGAGAC 64  
QY 991 GACGACATGGCGATGGTCTGGATTTCCATTGGCGACGAGGACGACGACGTTAGCTAT 1050  
Db 65 GACGACATGGCGATGGTCTGGATTTCCATTGGCGACGAGGAGGACGACGACGTTAGCTAT 124  
QY 1051 AGCTCTGTGGATGATGTTGGCGCAGACTACCGGACTACCGGATGATGTTAAATAAATCTC 1110  
Db 125 AGCTCTGTGGATGATGTTGGCGCAGACTACCGGACTACCGGATGATGTTAAATAAATCTC 184  
QY 1111 AACAAATGACATACCGGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 1170  
Db 185 AACAAATGACATACCGGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 244  
QY 1171 ACGGACATGTCATCTCAGCCTCAAAATGATGACATGATGTTTCCGATGTTTACCAGCTAC 1230  
Db 245 ACGGACATGTCATCTCAGCCTCAAAATGATGACATGATGTTTCCGATGTTTACCAGCTAC 304  
QY 1231 AATGCCCAAAAAAGAACGAGGAGAAAAATCTCGCTCGATTGGCGATGTACGCAATGAG 1290  
Db 305 AATGCCCAAAAAAGAACGAGGAGAAAAATCTCGCTCGATTGGCGATGTACGCAATGAG 364  
QY 1291 GAGCAGATATTCAGGGAATACACAGAGCTTCAGGAAAGTATCAATGAGGCAACT 1350  
Db 365 GAGCAGATATTCAGGGAATACACAGAGCTTCAGGAAAGTATCAATGAGGCAACT 424  
QY 1351 TCCAAAGAGAGCCCTGCAACCACTTCCACCGTCGAGAAATGTCATTCGCGCATTCGCCAC 1410  
Db 425 TCCAAAGAGAGCCCTGCAACCACTTCCACCGTCGAGAAATGTCATTCGCGCATTCGCCAC 484  
QY 1411 CTCCTAGTCCGCAA-----AGCAGATCCGAGGACTCGAGGCCAGCA 1452  
Db 485 CTCCTAGTCCGCAAAGGTGAATCTCTTTTCAGCCAGATCCGAGGACTCGAGGCCAGCA 544  
QY 1453 GCCCATTTCCACTTTGAGCAGCGCGGCTCACCAGGAAGTATGGGCTACCATGGAGAT 1512  
Db 545 GCCCATTTCCACTTTGAGCAGCGCGGCTCACCAGGAAGTATGGGCTACCATGGAGAT 604  
QY 1513 ATGTACATAGGAATGATTAACGAGAGAAATCTTTATCAGGAGCACTTTTCAACGCGCGAT 1572

Db 605 ATGTACATAGGAATGATAACGAGAGAAAATCTTATCAGGGACACTTTCAAACGCGCAT 664  
Qy 1573 GCGGCTTTGACGGTGACCAATACAGGCTTATATTACGTATACGGCCAGATATGCTACAAAC 1632  
Db 665 GCGGCTTTGACGGTGACCAATACAGGCTTATATTACGTATACGGCCAGATATGCTACAAAC 724  
Qy 1633 AACTGCGACGACCAAGACCGATTTATCGTCTTTCAAGGAGACACTCCATTCCTGCAAGTGC 1692  
Db 725 AACTGCGACGACCAAGACCGATTTATCGTCTTTCAAGGAGACACTCCATTCCTGCAAGTGC 784  
Qy 1693 TTGAACACGGTGCCCAACCAATGSCACATAGGTGCAACCTGSCACACAGATGCTGTG 1752  
Db 785 TTGAACACGGTGCCCAACCAATGSCACATAGGTGCAACCTGSCACACAGATGCTGTG 844  
Qy 1753 ATCCACCTGGAAACGAAACGAGAGGATCCATCTGAAGGACATTTCAACACGATCGCAATGCA 1812  
Db 845 ATCCACCTGGAAACGAAACGAGAGGATCCATCTGAAGGACATTTCAACACGATCGCAATGCA 904  
Qy 1813 GTTCTGCGGAGGAAACCAACGAAAGTACTTTTGGCATCTTTCAAGGTGTAATTTGGAGAG 1872  
Db 905 GTTCTGCGGAGGAAACCAACGAAAGTACTTTTGGCATCTTTCAAGGTGTAATTTGGAGAG 964  
Qy 1873 ATTATCCCGGTGAGAAATGGAATACCAATTTAAGCTTTTGTCCCGGACTGCTCGTG 1932  
Db 965 ATTATCCCGGTGAGAAATGGAATACCAATTTAAGCTTTTGTCCCGGACTGCTCGTG 1024  
Qy 1933 AATGCGATTATCGGCACGCGTGAATCCATTTAGTTCGTAGTACCTAGTCTTAGTCACTCCA 1992  
Db 1025 AATGCGATTATCGGCACGCGTGAATCCATTTAGTTCGTAGTACCTAGTCTTAGTCACTCAT 1084  
Qy 1993 AACCTAATCTCAATCGGAATCGTGATCTACTGATCTAGTCTAGTCTAGTCTAGTCTAGTCT 2052  
Db 1085 AACCTAATCTCAATCGGAATCGTGATCTACTGATCTAGTCTAGTCTAGTCTAGTCTAGTCT 1144  
Qy 2053 TTTATTTTGTATATCTCGTTCGACTCTAAAGAGTGAATATAATATATAGTCTATTTAA 2112  
Db 1145 TTTATTTTGTATATCTCGTTCGACTCTAAAGAGTGAATATAATATATATAGTCTATTTAA 1204  
Qy 2113 AAAAAAAAAAAAAAAAAA 2129  
Db 1205 AAAAAAAAAAAAAAAAAA 1221

RESULT 7  
AC012854  
LOCUS AC012854 55359 bp DNA linear HTG 03-NOV-1999  
DEFINITION Drosophila melanogaster, \*\*\* SEQUENCING IN PROGRESS \*\*\*.  
AC012854  
VERSION HTG; HTGS\_PHASE2.  
KEYWORDS Drosophila melanogaster (fruit fly)  
SOURCE Drosophila melanogaster  
ORGANISM Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.  
1 (bases 1 to 55359)  
Adams, M. and Venter, J.C.  
Direct Submission  
Submitted (03-NOV-1999) Celera Genomics, 45 West Gude Drive, Rockville, MD, USA  
This sequence was identified as CDM:10210834 by the submitter.  
For further information on this sequence you may e-mail to fly@celera.com.  
\* NOTE: This is a 'working draft' sequence.  
\* This sequence will be replaced  
\* by the finished sequence as soon as it is available and  
\* the accession number will be preserved.  
Location/Qualifiers  
1..55359  
/organism="Drosophila melanogaster"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:7227"

ORIGIN  
Query Match 37.2%; Score 798.8; DB 2; Length 55359;  
Best Local Similarity 99.8%; Pred. No. 5.9e-171;  
Matches 800; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
Qy 5 CGAGCGAAGCGAGCGTTTAAAGTGAAGAAACCGGTAAATCAGAGATCCCAAGCAAG 64  
Db 28568 CGAGCGAAGCGAGCGTTTAAAGTGAAGAAACCGGTAAATCAGAGATCCCAAGCAAG 28627  
Qy 65 CGCGTGGTGCATGATAGCGAAGAAAGAAAGCTATCCGTTTTCAGTTAACTTACTTACCAG 124  
Db 28628 CGCGTGGTGCATGATAGCGAAGAAAGAAAGCTATCCGTTTTCAGTTAACTTACTTACCAG 28687  
Qy 125 ATTGAATTTTCCCATCGGCAAAATTTACTAAATAATACATAAGTGAACCTCGTCCACTGTGT 184  
Db 28688 ATTGAATTTTCCCATCGGCAAAATTTACTAAATAATACATAAGTGAACCTCGTCCACTGTGT 28747  
Qy 185 GTTGTGTTTTTTTTTTTTTTTTTTTTTGGTTTTTGGCTGTGCTTTTATCGCAAAACGAATCGAT 244  
Db 28748 GTTGTGTTTTTTTTTTTTTTTTTTTTTGGTTTTTGGCTGTGCTTTTATCGCAAAACGAATCGAT 28807  
Qy 245 AAAAAGTGAAGAAATATCTTGAGAACTTTGTTTTCGGGCTTTTCTTTTGTCTAATTCGCCATC 304  
Db 28808 AAAAAGTGAAGAAATATCTTGAGAACTTTGTTTTCGGGCTTTTCTTTTGTCTAATTCGCCATC 28867  
Qy 305 GCGGAAGAGAAAAACAAGCAGTAGACAAAAACAAGTGTGTAATAACAATCTCGAAAGGGCA 364  
Db 28868 GCGGAAGAGAAAAACAAGCAGTAGACAAAAACAAGTGTGTAATAACAATCTCGAAAGGGCA 28927  
Qy 365 CCATCAGAGCCCGAGGGGTTTTTCTATATAGATGTGCGAGCTTATCATCTCATCTGCTGTC 424  
Db 28928 CCATCAGAGCCCGAGGGGTTTTTCTATATAGATGTGCGAGCTTATCATCTCATCTGCTGTC 28987  
Qy 425 TGTGAGTTGTTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 484  
Db 28988 TGTGAGTTGTTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 29047  
Qy 485 GTGCGACAAAGCTCGATTGGAAACAGCTGTGAGTGCCTTTGAGTGGTGGGCAAGATCG 544  
Db 29048 GTGCGACAAAGCTCGATTGGAAACAGCTGTGAGTGCCTTTGAGTGGTGGGCAAGATCG 29107  
Qy 545 TCATCATCATCATCGTCTGCTCATTTATCAACAGAAATCAGCATCAGCATCTGGAGGCCCGGT 604  
Db 29108 TCATCATCATCATCGTCTGCTCATTTATCAACAGAAATCAGCATCAGCATCTGGAGGCCCGGT 29167  
Qy 605 TGCTCTAAGATCCCGAGTGTTCATCAATATGACTGCGAGACCTCAAGCCGTTTATAA 664  
Db 29168 TGCTCTAAGATCCCGAGTGTTCATCAATATGACTGCGAGACCTCAAGCCGTTTATAA 29227  
Qy 665 CGCCAAGAGTGCACAGATGATGTTTTTCGGGCAAGCGACAGCACCGCGACCGGCC 724  
Db 29228 CGCCAAGAGTGCACAGATGATGTTTTTCGGGCAAGCGACAGCACCGCGACCGGCC 29287  
Qy 725 AGCGACACCGCGCAGCTGATCCCGTGGTTTTTGGGGTTTATCGTCTGGGGCTGGTCTG 784  
Db 29288 AGCGACACCGCGCAGCTGATCCCGTGGTTTTTGGGGTTTATCGTCTGGGGCTGGTCTG 29347  
Qy 785 TTGCCATTCTCGCACTAACGAT 806  
Db 29348 TTGCCATTCTCGCACTAACGAT 29369  
RESULT 8  
AC005974/c  
LOCUS AC005974 158983 bp DNA linear INV 19-NOV-1998  
DEFINITION Drosophila melanogaster, chromosome 2R, region 46E1-46F6, P1 clones DS05033, DS01913, and DS05181, complete sequence.  
AC005974 AC005433 AC005436 AC005469  
AC005974.1 GI:3893034  
KEYWORDS HTG.  
SOURCE Drosophila melanogaster (fruit fly)  
ORGANISM Drosophila melanogaster

REFERENCE AUTHORS	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.	
	1 (bases 1 to 158983)	
TITLE JOURNAL REFERENCE AUTHORS	Celnikier,S.E., Agbayani,A., Arcaina,T.T., Baxter,E., Blazej,R.G., Butenhoff,C., Champe,M., Chavez,C., Chew,M., Ciesiolka,L., Doyle,C.M., Farfan,D.E., Galle,R., George,R.A., Harris,N.L., Hoskins,R.A., Houston,K.A., Hummasti,S.R., Karra,K., Kearney,L., Kim,E., Lee,B., Lewis,S., Li,P., Lomotan,M.A., Mazda,P., Moshrefi,A.R., Moshrefi,M., Nixon,K., Pacleb,J.M., Park,S., Pfeiffer,B., Poon,L., Punch,E., Sequeira,A., Sethi,H., Snir,E., Svirkas,R.R., Twomey,B., Wan,K.H., Weinburg,T., Zhang,R., Zieran,L.L. and Rubin,G.M.	
	Sequencing of Drosophila chromosome 2R, region 46E1-46F6 Unpublished (1998)	
TITLE JOURNAL REFERENCE AUTHORS	2 (bases 1 to 158983)	
	Celnikier,S.E., Agbayani,A., Arcaina,T.T., Baxter,E., Blazej,R.G., Butenhoff,C., Champe,M., Chavez,C., Chew,M., Ciesiolka,L., Doyle,C.M., Farfan,D.E., Galle,R., George,R.A., Harris,N.L., Hoskins,R.A., Houston,K.A., Hummasti,S.R., Karra,K., Kearney,L., Kim,E., Lee,B., Lewis,S., Li,P., Lomotan,M.A., Mazda,P., Moshrefi,A.R., Moshrefi,M., Nixon,K., Pacleb,J.M., Park,S., Pfeiffer,B., Poon,L., Punch,E., Sequeira,A., Sethi,H., Snir,E., Svirkas,R.R., Twomey,B., Wan,K.H., Weinburg,T., Zhang,R., Zieran,L.L. and Rubin,G.M.	
TITLE JOURNAL	Direct Submission	
	Submitted (19-NOV-1998) Berkeley Drosophila Genome Project, MS 64-121, Lawrence Berkeley National Laboratory, One Cyclotron Road, Berkeley, CA 94720, US	
COMMENT	Sequence submitted by:	
	Berkeley Drosophila Genome Project Lawrence Berkeley National Laboratory, MS 64-121 Berkeley, CA 94720	
FEATURES source	For further information about this sequence, including its location and relationship to other sequences, please visit our sequence archive Web site ( <a href="http://www.fruitfly.org/sequence/">http://www.fruitfly.org/sequence/</a> ) or send email to <a href="mailto:bdgp@fruitfly.berkeley.edu">bdgp@fruitfly.berkeley.edu</a> .	
	PI library locations: 53-41, 20-89, 54-93. Location/Qualifiers 1..158983 /organism="Drosophila melanogaster" /mol_type="genomic DNA" /strain="y2; cn bw sp" /db_xref="caxon:7227" /chromosome="2R" /map="46E1-46F6" /clone="P1s D305033 (D347), D301913 (D350), and D305181 (D384)" /clone_lib="P1 library, partial Sau3A in pAd10acBII" /note="These P1s were completed as a project. D305181 (D384) is a partial bridge with its distal neighbor extending from bp 1 to P1 end at bp 57,624. D301913 (D350) extends from P1 end at 7,924 to P1 end at 92,484. D305033 (D347) extends from P1 end at bp 80,881 to P1 end at bp 158,983. Between bps 53,139 and 53,140, Tn10 was excised from D305181 along with an associated 9 bp duplication."	
ORIGIN	Query Match 37.2%; Score 798.8; DB 3; Length 158983; Best Local Similarity 99.8%; Pred. No. 7.9e-171; Matches 800; Conservative 0; Mismatches 2; Indels 0; Gaps 0;	
	5 CGAGCGCAACGGAGCTTTAAAGTCAGAGAAAGAAACCGGTAATTCAGAGATCCCAAGCAAG 64	
Db	134656	CGAGCGCAACGGAGCTTTAAAGTCAGAGAAAGAAACCGGTAATTCAGAGATCCCAAGCAAG 134597
Qy	65	CGCGTGGCGTCATGATAGCGAAGAAAAGCTATCGTTTCAGTTAACTACTTACCAAG 124
	Db	134596 CGCGTGGCGTCATGATAGCGAAGAAAAGCTATCGTTTCAGTTAACTACTTACCAAG 134537
Qy	125	ATTGAATTTCCGCAATCGGGCAATTTACTAAATAATACATAGTGCAACTCGTCCACTGTGT 184
	Db	134536 ATTGAATTTCCGCAATCGGGCAATTTACTAAATAATACATAGTGCAACTCGTCCACTGTGT 134477

Qy	185	GTTCGTGTTTTTTTTTTTTTTTTTTTTTTTGGTTTTTCGCTGTGCTTTTATCGCAAAACAAGACTGAT 244
Db	134476	GTTCGTGTTTTTTTTTTTTTTTTTTTTTTTGGTTTTTCGCTGTGCTTTTATCGCAAAACAAGACTGAT 134417
Qy	245	AAAACCTAGAAAATATCTTGAGAACTTGTTTTTCGCGCTTTTCTTTTTCGTAATTCGCGATC 304
	Db	134416 AAAACCTAGAAAATATCTTGAGAACTTGTTTTTCGCGCTTTTCTTTTTCGTAATTCGCGATC 134357
Qy	305	CGCGAAGAGAAAAACAAGCAGTAGACAAAACRAAGTGTGGTAATACAACTCTGAAAAGGGCA 364
	Db	134356 CGCGAAGAGAAAAACAAGCAGTAGACAAAACRAAGTGTGGTAATACAACTCTGAAAAGGGCA 134297
Qy	365	CCATCAGCAGCCCGAGGGTTTTATCTATATAGATGTGCGAGCTTATCATCTCATGTGTC 424
	Db	134296 CCATCAGCAGCCCGAGGGTTTTATCTATATAGATGTGCGAGCTTATCATCTCATGTGTC 134237
Qy	425	TGTGAGGTGTTCTGTGTGCTGTGTAGTATCTTAAATACATAGAGTGTGTTTCATATAAA 484
	Db	134236 TGTGAGGTGTTCTGTGTGCTGTGTAGTATCTTAAATACATAGAGTGTGTTTCATATAAA 134177
Qy	485	GTGCGCAAGAGCTCGAATTGGAACACGCTGTCGAGTGCCTTGAGTGGGTGGCAGATCG 544
	Db	134176 GTGCGCAAGAGCTCGAATTGGAACACGCTGTCGAGTGCCTTGAGTGGGTGGCAGATCG 134117
Qy	545	TCATCATCATCATCTGTCGTCATTTATCAACAGAAATCAGCATCAGCATCTGGAGGCCCGCGT 604
	Db	134116 TCATCATCATCATCTGTCGTCATTTATCAACAGAAATCAGCATCAGCATCTGGAGGCCCGCGT 134057
Qy	605	TGCTCTAAGATCCCGAGTGTTCATCAATTTATGACCTGCGAGACCCCTCAAGCGTTTATAA 664
	Db	134056 TGCTCTAAGATCCCGAGTGTTCATCAATTTATGACCTGCGAGACCCCTCAAGCGTTTATAA 133997
Qy	665	CGCCAAACGAGTGCCTGATGATGTTTTTCCGCGCAAAAGCCAGCAGCGCGCCGCC 724
	Db	133996 CGCCAAACGAGTGCCTGATGATGTTTTTCCGCGCAAAAGCCAGCAGCGCGCCGCC 133937
Qy	725	ACGCGAGCACCCCGCAGCTGATCCCGCTGGTTTGGGGTTCATCGGTCTGGGGCTGGTGC 784
	Db	133936 ACGCGAGCACCCCGCAGCTGATCCCGCTGGTTTGGGGTTCATCGGTCTGGGGCTGGTGC 133877
Qy	785	TTGCCATTCTCGCACTAACGAT 806
	Db	133876 TTGCCATTCTCGCACTAACGAT 133855
RESULT 9		
AC007414		
LOCUS		
DEFINITION		
Drosophila melanogaster, chromosome 2R, region 46C-46D, BAC clone BACR06J01, complete sequence.		
AC007414		
AC007414.6 GI:15451491		
HTG.		
SOURCE		
Drosophila melanogaster (fruit fly)		
Drosophila melanogaster		
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.		
1 (bases 1 to 172904)		
REFERENCE		
AUTHORS		
Celnikier,S.E., Adams,M.D., Kronmiller,B., Tyler,D., Wan,K.H., Holt,R.A., Evans,C.A., Gocayne,J.D., Ananides,P.G., Brandon,R.C., Rogers,Y., An,H., Baldwin,D., Banzon,J., Beeson,K.Y., Busam,D.A., Carlson,J.W., Center,A., Champe,M., Davenport,L.B., Dietz,S.M., Dodson,K., Dorsett,V., Doup,E.E., Doyle,C., Dresnek,D., Farfan,D., Ferreira,S., Frise,E., Galle,R.F., Garg,N.S., George,R.A., Gonzalez,M., Houck,J., Hoskins,R.A., Hostin,D., Howland,T.J., Ibegwam,Y., Jalali,M., Kruse,D., Li,P., Mattei,B., Moshrefi,A., McIntosh,T.C., Moy,M., Murphy,B., Nelson,C., Nelson,K.A., Nunoo,J., Pacleb,J., Paragas,V., Park,S., Patel,S., Pfeiffer,B., Phouanavong,S., Pittman,G.S., Puri,V., Richards,S., Scheeler,F., Stapleton,M., Strong,R., Svirkas,R., Tector,C., Williams,S.M., Zaveri,J.S., Smith,H.O., Rubin,G.M. and Venter,J.C.		







64-121, Lawrence Berkeley National Laboratory, One Cyclotron Road, Berkeley, CA 94720, US  
Sequence submitted by:  
Berkeley Drosophila Genome Project  
Lawrence Berkeley National Laboratory, MS 64-121  
Berkeley, CA 94720

This sequence was assembled using end sequences from a whole genome shotgun and from subclones of this BAC and its neighboring clones. For further information about this sequence, including its location and relationship to other sequences, please visit our sequence archive web site (<http://www.fruitfly.org/sequence/>) or send email to [bdg@fruitfly.berkeley.edu](mailto:bdg@fruitfly.berkeley.edu).  
Location/Qualifiers

## FEATURES

Source

1. .189620  
/organism="Drosophila melanogaster"  
/mol\_type="genomic DNA"  
/strain="y; cn bw sp"  
/db\_xref="taxon:7227"  
/chromosome="2R"  
/map="46D-47A"  
/clone="BACR09J12 (D1345)"  
/clone.lib="RPCL-98 (Roswell Park Cancer Institute  
Drosophila melanogaster BAC library, partial EcoRI in  
pBACE3.6)"

## ORIGIN

Query Match 37.2%; Score 798.8; DB 3; Length 189620;  
Best Local Similarity 99.8%; Pred. No. 8.3e-171;  
Matches 800; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
  
QY 5 CGAGCGCAACGGACGTTTAAAGTGAGAAAAGAAACCGGTAAATCAGAGATCCCAAGCAAG 64  
DB 169512 CGACGCAAGCGGACGTTTAAAGTGAGAAAAGAAACCGGTAAATCAGAGATCCCAAGCAAG 169453  
  
QY 65 CGCGTGGCTGCATGATACGAGAAAAGAAAGCTATCGTTTCAGTTAACTACTTACCAAG 124  
DB 169452 CGCGTGGCTGCATGATACGAGAAAAGAAAGCTATCGTTTCAGTTAACTACTTACCAAG 169393  
  
QY 125 ATTGAATTTCCGCAATCGGGCAAAATACATAAGTGAACCTCGTCCACTGTGT 184  
DB 169392 ATTGAATTTCCGCAATCGGGCAAAATACATAAGTGAACCTCGTCCACTGTGT 169333  
  
QY 185 GTTGTTTTTTTTTTTTTTTTTTTTTTCGTTTCGCTGTGCTTTATCGCAAAACAAGACTGAT 244  
DB 169332 GTTGTTTTTTTTTTTTTTTTTTTTTTCGTTTCGCTGTGCTTTATCGCAAAACAAGACTGAT 169273  
  
QY 245 AAAAAGTGAATATCTTGAGAACTGTTTTTCGCGCTTTCTTTCTAATTCGCGATC 304  
DB 169272 AAAAAGTGAATATCTTGAGAACTGTTTTTCGCGCTTTCTTTCTAATTCGCGATC 169213  
  
QY 305 GCGGAAGAGAAAAACAAGCAGTAGACAAAACAAGTGTGTAATACATCTGAAAAGGGCA 364  
DB 169212 GCGGAAGAGAAAAACAAGCAGTAGACAAAACAAGTGTGTAATACATCTGAAAAGGGCA 169153  
  
QY 365 CCATCAGAGCCCGAGGGGTTTATCTATATAGATGTCGAGCTTATCATCTCATGCTGTC 424  
DB 169152 CCATCAGAGCCCGAGGGGTTTATCTATATAGATGTCGAGCTTATCATCTCATGCTGTC 169093  
  
QY 425 TGTGAGGTGTTGTCGTGCTCGTGTAGTATCTTAATACATAGAGTGTTTCATATAA 484  
DB 169092 TGTGAGGTGTTGTCGTGCTCGTGTAGTATCTTAATACATAGAGTGTTTCATATAA 169033  
  
QY 485 GTGCGAACAAGCTCGATTGAAAACAGCTGTGAGTGCCCTTGTAGTGGGTGGGCAAGATCG 544  
DB 169032 GTGCGAACAAGCTCGATTGAAAACAGCTGTGAGTGCCCTTGTAGTGGGTGGGCAAGATCG 168973  
  
QY 545 TCATCATCATCGTCTGTCATTATCAAGATCAGCATCAGCATCTGAGGCCCGCGGT 604  
DB 168972 TCATCATCATCGTCTGTCATTATCAAGATCAGCATCAGCATCTGAGGCCCGCGGT 168913  
  
QY 605 TGCTCTAAGATCCCGAGTGTTTCATCAATATGACTGCGAGACCCCTCAAGCCGTTTATAA 664  
DB 168912 TGCTCTAAGATCCCGAGTGTTTCATCAATATGACTGCGAGACCCCTCAAGCCGTTTATAA 168853

## COMMENT

QY 665 CGCCAAACGAGTGCACACGATGATGGTTTTCGGCCAAAGCAGCAGCAGCGCCGCC 724  
DB 168852 CGCCAAACGAGTGCACACGATGATGGTTTTCGGCCAAAGCAGCAGCAGCGCCGCC 168793  
  
QY 725 AGGAGCGACCGCCGAGCTGATCCCTGGTTTGGGTTTCATCGTCTGGGCTGGTGC 784  
DB 168792 AGGAGCGACCGCCGAGCTGATCCCTGGTTTGGGTTTCATCGTCTGGGCTGGTGC 168733  
  
QY 785 TTGCGATTCTCGCACTAACGAT 806  
DB 168732 TTGCGATTCTCGCACTAACGAT 168711

## RESULT 11

AE003831

## LOCUS

AE003831 Drosophila melanogaster chromosome 2R, section 19 of 74 of the complete sequence. 275390 bp DNA linear INV 14-FEB-2003

## ACCESSION

AE003831 AE002787 AE013599

## VERSION

AE003831.3 GI:21645483

## KEYWORDS

Drosophila melanogaster (fruit fly)

## SOURCE

Drosophila melanogaster

## ORGANISM

Drosophila melanogaster

Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

## REFERENCE

1 (bases 1 to 275390)

## AUTHORS

Adams,M.D., Celisner,S.E., Holt,R.A., Evans,C.A., Gocayne,J.D., Amanatides,D., Scherer,S.E., Li,P.W., Hoskins,R.A., Galle,R.F., George,R.A., Lewis,S.E., Richards,S., Ashburner,M., Henderson,S.N., Sutton,G.G., Wortman,J.R., Yandell,M.D., Zhang,Q., Chen,L.X., Brandon,R.C., Rogers,Y.H., Blazer,J.R.G., Champe,M., Pfeiffer,B.D., Wan,K.H., Doyle,C., Baxter,E.G., Helt,G., Nelson,C.R., Gabor,G.L., Ahril,J.F., Agbayani,A., An,H.J., Andrews-Pfannkoch,C., Baldwin,D., Ballew,R.M., Basu,A., Baxendale,J., Bayraktaroglu,L., Beasley,E.M., Beeson,K.Y., Bonin,P.V., Borman,B.P., Bhandari,D., Bolshakov,S., Borkova,D., Botchan,M.R., Bouck,J., Brokstein,P., Brotter,P., Burdick,C., Busam,D.A., Butler,H., Cadieu,E., Center,A., Chandra,I., Cherry,J.M., Cawley,S., Dahlke,C., Davenport,L.B., Davies,P., de Pablos,B., Delcher,A., Deng,Z., Mays,A.D., Dew,I., Dietz,S.M., Dodson,K., Doup,L.E., Downes,M., Dugan-Rocha,S., Dunkov,B.C., Dunn,P., Durbin,K.J., Evangelista,C.C., Ferraz,C., Ferreira,S., Fleischmann,W., Fosler,C., Gabriellian,A.E., Garu,N.S., Gelbart,W.M., Glasser,K., Glodek,A., Gong,F., Gorrell,J.H., Gu,Z., Guan,P., Harris,M., Harris,N.L., Harvey,D., Heiman,T.J., Hernandez,J.R., Houck,J., Hostin,D., Houston,K.A., Howland,T.J., Wei,M.H., Ibegwam,C., Jalali,M., Kalush,F., Karpen,G.H., Ke,Z., Kennison,J.A., Ketchum,K.A., Kimmel,B.E., Kodira,C.D., Kraft,C., Kravitz,S., Kulp,D., Lai,Z., Lasko,P., Lei,Y., Levitsky,A.A., Li,J., Li,Z., Liang,Y., Lin,X., Liu,X., Ma,J., McIninch,T.C., McLeod,M.P., McPherson,D., Merkulov,G., Milshina,N.V., Mobarry,C., Morris,J., Moshrefi,A., Mount,S.M., Moy,M., Murphy,B., Murphy,L., Muzny,D.M., Nelson,D.L., Nelson,D.R., Nelson,K.A., Nixon,K., Nusskern,D.R., Pacle,J.M., Palazzolo,M., Pittman,G.S., Pan,S., Pollard,J., Puri,V., Reese,M.G., Reinert,K., Remington,K., Saunders,R.D., Scheeler,F., Shen,H., Shue,B.C., Siden-Kiamos,I., Simpson,M., Skupski,M.P., Smith,T., Spier,E., Spradling,A.C., Stapleton,M., Strong,R., Sun,B., Svirskas,R., Tector,C., Turner,R., Venter,E., Wang,A.H., Wang,X., Wang,Z.Y., Wassarman,D.A., Weinstock,G.M., Weissbach,J., Williams,S.M., Woodage,T., Worley,K.C., Wu,D., Yang,S., Yao,Q.A., Ye,J., Yeh,R.F., Zaveri,J.S., Zhan,M., Zhang,G., Zhao,Q., Zheng,L., Zheng,X.H., Zhong,F.N., Zhong,W., Zhou,X., Zhu,S., Zhu,X., Smith,H.O., Gibbs,R.A., Myers,E.W., Rubin,G.M. and Venter,J.C.

## TITLE

The genome sequence of Drosophila melanogaster

## JOURNAL

Science 287 (5461), 2185-2195 (2000)

## MEDLINE

20196006

## PUBMED

10731132

## REFERENCE

2 (bases 1 to 275390)

## AUTHORS

Celisner,S.E., Adams,M.D., Kronmiller,B., Wan,K.H., Holt,R.A., Evans,C.A., Gocayne,J.D., Amanatides,P.G., Brandon,R.C., Rogers,Y., Banzon,J., An,H., Baldwin,D., Banzon,J., Beeson,K.Y., Busam,D.A.,

Carlson,J.W., Center,A., Champe,M., Davenport,L.B., Dietz,S.M., Dodson,K., Dorsett,V., Doup,L.E., Doyle,C., Dresnek,D., Farfan,D., Ferreira,S., Friese,E., Galle,R.F., Garg,N.S., George,R.A., Gonzalez,M., Houck,J., Hoskins,R.A., Hostin,D., Howland,T.J., Ibegwam,C., Jalali,M., Kruse,D., Li,P., Matzel,B., Moshrefi,A., McIntosh,F.C., Moy,M., Murphy,B., Nelson,C., Nelson,K.A., Nunoo,J., Pacle,J., Paragas,V., Park,S., Patel,S., Pfeiffer,B., Phouanavong,S., Pittman,G.S., Puri,V., Richards,S., Scheeler,F., Stapleton,M., Strong,R., Svirskas,R., Tector,C., Tyler,D., Williams,S.M., Zaveri,J.S., Smith,H.O., Venter,J.C. and Rubin,G.M. Sequencing of *Drosophila melanogaster* genome Unpublished  
3 (bases 1 to 275390)  
Miera,S., Crosby,M.A., Matthews,B.B., Bayraktaroglu,I., Campbell,K., Hradecky,P., Huang,Y., Kaminker,J.S., Prochnik,S.E., Smith,C.D., Tupy,J.L., Bergman,C.M., Berman,B.P., Carlson,J.W., Celinker,S.E., Clamp,M.E., Drysdale,R.A., Emmert,D., Friese,E., de Grey,A.D.N.J., Harris,N.L., Kronmiller,B., Marshall,B., Milburn,G.H., Richter,J., Russo,S., Searle,S.M.J., Smith,E., Shu,S., Smutniak,F., Whitfield,E.J., Ashburner,M., Gelbart,W.M., Rubin,G.M., Mungall,C.J. and Lewis,S.E. Annotation of *Drosophila melanogaster* genome Unpublished  
4 (bases 1 to 275390)  
Adams,M.D., Celinker,S.E., Gibbs,R.A., Rubin,G.M. and Venter,C.J. Direct Submision  
Submitted (21-MAR-2000) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA  
5 (bases 1 to 275390)  
FlyBase  
Direct Submision  
Submitted (06-SEP-2002) University of California Berkeley, 539 Life Sciences Addition, Berkeley, CA 94720, USA  
6 (bases 1 to 275390)  
FlyBase  
Direct Submision  
Submitted (23-JAN-2003) University of California Berkeley, 539 Life Sciences Addition, Berkeley, CA 94720, USA  
On Jul 1, 2002 this sequence version replaced gi:10727672.  
Location/Qualifiers  
1. .275390  
/organism="Drosophila melanogaster"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:7227"  
/chromosome="2R"  
/note="genotype: y[1]; cn[1] bw[1] ep[1]; Rh6[1]"  
197. .2954  
/locus\_tag="CG1671"  
/note="last curated on Mon Feb 11 17:53:43 PST 2002"  
/map="46B7-46B9"  
/db\_xref="FLYBASE:FBgn0033454"  
join(197. .286,346. .2954)  
/locus\_tag="CG1671"  
/product="CG1671-RA"  
/db\_xref="FLYBASE:FBgn0033454"  
join(261. .286,346. .2683)  
/locus\_tag="CG1671"  
/note="CG1671 gene product"  
/codon\_start=1  
/product="CG1671-PA"  
/protein\_id="AAF58088.1"  
/db\_xref="GI:7303841"  
/db\_xref="FLYBASE:FBgn0033454"  
/translation="MLAGISEKSYAAEARVTFNYAGGDIAMSPNGHHLYCLINGSANVQ VQDTGSLNSYSSPSADNKRPAELNDIDVEDTITCGLSQELHVTARSGLLRL WLDTGKLVKQAQHGKPVIRVFSPCGRLICTSGGADATLRWDYSNNSCIGALKXD PPGAWLLVHPNVLRKVIYAGGSDNTVYAWNYETKLVMRCHLSQVTGLSPKNA SDCNEIVTSRDKVMIWVQEQESKQLKVLPLYDELEGVYVADEQLLIIVASGSGK LQQVDPKSWKIRDSDFQISRLHCKSELQALVTTQENILVYDTELEPFIKHLV GYNDEILDMCFMGHDEYLAVATNSKHFKLYDTEHDMNCKLIVGHSDTVMSLAQNL LISVGKDCSVRLMLQHDKDCLEALTQQANCHTSTTGCVAMTHNGATGFSVSDGS

MKVQLVRSKEDRNSYSFNLRYYAALSHDKVEVNCVAPANNKLIATASODTKAKWLAE SNTLQGLGTRGVMSRPSPDQIVLTSSDCTLRISISNFSCTIKRFDQECTILR AEFIDGKGTIISAASDGLLKNWIKNTQISLDENDRVYALAVSARSNRFYTGGA DSKLIIPGDVQVTRNEALDKROANLQEOTLHSLHAQQLHKAFVLALNLDPKPAS FDIINPVRKDEPQQLVDQNVQRAVLLQHVKAWTNRSRHSQAGNMILKYLIGD ALIDPKARFYNGNMVEVTFYVQRHFKRVTELNKDLAFLEFVYKCM"  
complement (2844. .13679)  
/gene="BcdNA.GH06348"  
/locus\_tag="CG1516"  
/note="last curated on Tue Mar 12 18:58:27 PST 2002"  
/map="46B9-46B12"  
/db\_xref="FLYBASE:FBgn0027580"  
complement (join(2844. .3521,3590. .4619,4684. .5341, 6647. .8156,8220. .8346,13587. .13679))  
/gene="BcdNA.GH06348"  
/locus\_tag="CG1516"  
/product="CG1516-RE"  
/db\_xref="FLYBASE:FBgn0027580"  
complement (join(2844. .3521,3590. .4619,4684. .5341, 5819. .5866,6647. .8156,8220. .8346,13587. .13679))  
/gene="BcdNA.GH06348"  
/locus\_tag="CG1516"  
/product="CG1516-R1"  
/db\_xref="FLYBASE:FBgn0027580"  
complement (join(2844. .3521,3590. .4619,4684. .5341, 6647. .8156,8220. .8346,12506. .12588,12661. .12893))  
/gene="BcdNA.GH06348"  
/locus\_tag="CG1516"  
/product="CG1516-RD"  
/db\_xref="FLYBASE:FBgn0027580"  
complement (join(2844. .3521,3590. .4619,4684. .5341, 6647. .8156,8220. .8346,12506. .12893))  
/gene="BcdNA.GH06348"  
/locus\_tag="CG1516"  
/product="CG1516-RA"  
/db\_xref="FLYBASE:FBgn0027580"  
complement (join(2844. .3521,3590. .4619,4684. .5341, 5819. .5866,6647. .8156,8220. .8346,12506. .12893))  
/gene="BcdNA.GH06348"  
/locus\_tag="CG1516"  
/product="CG1516-RJ"  
/db\_xref="FLYBASE:FBgn0027580"  
complement (join(2844. .3521,3590. .4619,4684. .5341, 6647. .8156,8220. .8346,8412. .8546))  
/gene="BcdNA.GH06348"  
/locus\_tag="CG1516"  
/product="CG1516-RB"  
/db\_xref="FLYBASE:FBgn0027580"  
complement (join(2844. .3521,3590. .4619,4684. .5341, 5819. .5866,6647. .8156,8220. .8346,8412. .8546))  
/gene="BcdNA.GH06348"  
/locus\_tag="CG1516"  
/product="CG1516-RK"  
/db\_xref="FLYBASE:FBgn0027580"  
complement (join(2844. .3521,3590. .4619,4684. .5341, 6647. .8156,8220. .8544))  
/gene="BcdNA.GH06348"  
/locus\_tag="CG1516"  
/product="CG1516-RG"  
/db\_xref="FLYBASE:FBgn0027580"  
complement (join(2844. .3521,3590. .4619,4684. .5341, 5819. .5866,6647. .8156,8220. .8544))  
/gene="BcdNA.GH06348"  
/locus\_tag="CG1516"  
/product="CG1516-RL"  
/db\_xref="FLYBASE:FBgn0027580"  
complement (join(3273. .3521,3590. .4619,4684. .5341, 6647. .8156,8220. .8318))  
/gene="BcdNA.GH06348"  
/locus\_tag="CG1516"

Query Match 37.2%; Score 798.8; DB 3; Length 275390;  
Best Local Similarity 99.8; Pred. No. 9.3e-171;





JOURNAL Patent: WO 0300899-A 5263 03-JAN-2003;  
Syngenta Participations AG (CH)  
FEATURES Location/Qualifiers  
source 1..2000  
/organism="Oryza sativa"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:4530"

Qy 1610 TATA 1613  
Db 905 TCWW 908

Search completed: September 15, 2004, 00:00:18  
Job time : 13370 secs

## ORIGIN

Query Match	3.1%	Score 66.4	DB 6	Length 2000
Best Local Similarity	10.0%	Pred. No. 0.00046		
Matches	90	Conservative 415	Mismatched 391	Indels 8
	Gaps			
Qy	710	GCACGGCAGCCCGCAGCGAGCAGCCCGCCAGCTGATCCCCTGGTTTGGGGTTCATCG	769	
Db	13	SSWRGRRMYWAGMMSCHARGSSRMSRKWGSYKRYKSSCGKCMKMTTTRRSKWYSASSA	72	
Qy	770	GTCTGGGCTGGTCTGCTTGCCTATCTTCGCACCTAACGATCTGCAGACACCGGTGTATCGC	829	
Db	73	SGRTGSKWSSGSYSGKMKRYKRSKRWRGRGRGRMRMRWRMRRCARSRMRAGGS	132	
Qy	830	ATCTGGAACAAGAGCTGAAGAGCTGAAGCAGCTGTCGATAATCTCCAGCAGCGTTTGG	889	
Db	133	GRMMGGKRSYVWMCYARGGSCKRKKS KGSWGKTCRRGARGSGWSGAKYKSGSMS	192	
Qy	890	GCATAAACTATCTGACGAGTTCGACGAGTTCCAAAAGGAGTACGAGATGCCCTCATCG	949	
Db	193	KRMWSSCGRGGRSAYSIYGTSRKGYTKMYYSASRCNRYMTTYSYMWACSSYT	252	
Qy	950	ACTATCCAAAAGGTGATGGCTTCACGGATGAGGAGGACGACGATGGCGCATGGTC	1009	
Db	253	WCRSKRRSMWMMKMRWSSSYGWYSWYKWMCTAYKKSYSRWCYVWGGWRATR	312	
Qy	1010	TGGATTCCAATTGGCGACGACGAGCAGCAGCTAGCTATGACTCTGTGGATGATGTTG	1069	
Db	313	YWGGRYSRMAWMTKMYTYGKMGKRWGAGRMWMSRCWRSKACYVWRMRWMTTR	372	
Qy	1070	GCGCAGACTACGAGACTACACCCATATGTTAAATAAACTCAACATGACATACCGCA	1129	
Db	373	RRWAKKSRTSRKKRWKWKRYKMRGYSRMRSCRRMRWKRSGRAWKMGCRGM	432	
Qy	1130	CCAGCCCAACATCTGAGACCACTGCTGAGGCGGAGGCGAGACGACATGTCATCTCTCAG	1189	
Db	433	TCRMKSYGMRWKSWMAS KYKMSRMYRWKKKCSRTTMMWKTGRGMMGTWGRCRYK	492	
Qy	1190	CCTCAAATGATGACAAATGTTTCATGACTTACCAGCTACATGCCCCACAAAAGAACG	1249	
Db	493	RSWMKRCRRRWGRMRWKRKYMSAYTMYRCARKK---YSYSAARKCWVRGKG	548	
Qy	1250	AGGAGAAATCTCGCTCGATTCCCGATGTAGCAATGAGGACGAGCAATATTCAGGAA	1309	
Db	549	YYWAGMWKRYKMYWYKMMWYKRYKSKSWYKMSYASCMKSAKAKGKCKRSKMS	608	
Qy	1310	ATCACAGAGCTTCAGAAAAGTCATCCAAATGAGGCAACTCCAAAGAGAGCGCTTCAC	1369	
Db	609	AWSKMSRPSRKCKACSKRSSAKYAMMGMTSGSRMRWSYCYVWRKWSMKSTCTWM	668	
Qy	1370	CACCTTACCACCGTCGAGAAATGCAATCCCGCATCGCCACTCTTAGTCCGCAAGCCA	1429	
Db	669	YMSKYTYAKGYSWRYRAW-CMYMRWYYRYRSYMTYMAWYTSSTRMAMTGMKYS	727	
Qy	1430	GATCCGAGGACTCGAGCCAGCACCATTTCACATTGACGACGCGCGCTCACAAG	1489	
Db	728	RYWTSWYKYCKSKWYSWYYSWVAKTWKWRRYATRMWMMWYRYSKMTYTWCTMWG	787	
Qy	1490	GAGATGTGGCTACCATGGAGATATGTACATAGGAAATGATACGAGAGAACTTTATC	1549	
Db	788	YMWYWRTYMKRMYMYKCTKYVWYSATVYTGTAW---WMAKTKMRMGMTGAKTRGR	844	
Qy	1550	AGGGACACITTCAAACGCGCGATGGCGCTTTCAGCGGTGACCAATACAGGCCTATATTAG	1609	
Db	845	ARKARYWWTWCACTKRWMTKGKAKWATWMAKWRKYYSWMPAWTYYYKTRTRYK	904	

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 14, 2004, 14:52:20 ; Search time 1222 Seconds  
(without alignments)  
7467.373 Million cell updates/sec

Title: US-09-813-329-5  
Perfect score: 2148  
Sequence: 1 ggcacgagcgacgcgcgt.....aaaaaaaaaactcgag 2148

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : N\_Geneseq\_29Jan04.\*

1: Geneseqn1980s.\*

2: Geneseqn1990s.\*

3: Geneseqn2000s.\*

4: Geneseqn2001as.\*

5: Geneseqn2001bs.\*

6: Geneseqn2002s.\*

7: Geneseqn2003as.\*

8: Geneseqn2003bs.\*

9: Geneseqn2003cs.\*

10: Geneseqn2004s.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,

and is derived by analysis of the total score distribution.

SUMMARIES

Result No. Score Query Match Length DB ID Description

1	2148	100.0	2148	6	ABK11680	Abk11680 DNA encod
2	2112.8	98.4	2166	6	ABK11678	Abk11678 DNA encod
3	1142.8	53.2	1221	6	ABK11679	Abk11679 DNA encod
4	907.4	42.2	978	4	ABL21473	Ab121473 Drosophil
5	779.4	36.3	3324	4	ABL21472	Ab121472 Drosophil
6	66.8	3.1	390	2	AAQ21833	Aa21833 Randomisi
7	66.8	3.1	390	2	AAQ36859	Aa36859 PCR prime
8	66.8	3.1	390	5	AAQ76910	Aa76910 Sequence
9	66.8	3.1	390	6	AAI72775	Aa72775 Oligo #7
10	66.4	3.1	2000	7	ADA71938	Ada71938 Rice gene
11	55.8	2.6	3489	3	AAA30290	Aa30290 Kaposi's s
12	55.8	2.6	3489	4	AAQ82901	Aa82901 Nucleotid
13	55.8	2.6	3489	6	ABA93487	Ab93487 Kaposi's s
14	55.8	2.6	32207	2	AAV73805	Aa73805 KSHV LUR
15	55.8	2.6	137507	2	AAV19941	Aa19941 KSHV long
16	54.4	2.5	5306	6	ABL32510	Ab132510 Human imm
17	54.2	2.5	445	6	ABZ23375	Abz23375 Reverse c
18	54.2	2.5	445	6	ABZ23374	Abz23374 Nucleotid
19	54	2.5	15935	4	ABL06114	Ab106114 Drosophil
20	53.6	2.5	1915	7	ABZ42562	Abz42562 Human C-C
21	53.6	2.5	1958	3	AAA27006	Aa27006 Human cdn
22	53	2.5	234	5	ABV57409	Abv57409 Human pro
23	52.6	2.4	1898	6	ABQ54407	Abq54407 Human ova

24	52.4	2.4	479	8	ACH22127	Ach22127 Human adu
c	52.4	2.4	5858	6	ABL33312	Ab133312 Human imm
25	52.4	2.4	422	4	AAI87575	Aa187575 Human pol
26	52	2.4	471	5	ABV56603	Abv56603 Human pro
27	52	2.4	1109	4	ABL03311	Ab103311 Drosophil
28	52	2.4	3116	4	ABL03310	Ab103310 Drosophil
29	51.6	2.4	550	8	ACH43937	Ach43937 Human toe
30	51.6	2.4	1037	3	AAA26447	Aa26447 Human sec
31	51.6	2.4	1037	7	ADA56497	Ada56497 Gene enco
32	51.6	2.4	1037	9	ADC73848	Adc73848 Human eec
33	51.6	2.4	1037	9	ADD37737	Add37737 Human eec
34	51.6	2.4	1037	9	AAH35025	Aah35025 Human col
35	51.6	2.4	1309	4	ABX41811	Abx41811 Bovine ES
c	51.4	2.4	337	7	AAI87804	Aa187804 Human pol
36	51.4	2.4	3527	2	AAT72171	Aat72171 Alzhelmer
37	51.4	2.4	323	5	ABV60996	Abv60996 Human pro
38	51.2	2.4	546	4	AAH71167	Aah71167 Human cer
39	51.2	2.4	14006	6	ABL33958	Ab133958 Human imm
40	51.2	2.4	375	7	ABX49849	Abx49849 Bovine ES
c	51	2.4	527	4	AAI43436	Aa143436 Probe #12
41	51	2.4	2317	8	ACF35812	Acf35812 Human met
42	51	2.4	3510	5	AAS01369	Aas01369 Human TAN
43	51	2.4				
44	51	2.4				
45	51	2.4				

ALIGNMENTS

RESULT 1

ABK11680

ID ABK11680 standard; DNA; 2148 BP.

XX AC ABK11680;

XX DT 05-JUN-2002 (first entry)

XX DE DNA encoding tumour necrosis factor variant 2 (TNFv2).

XX KW Tumour necrosis factor; TNF; pesticide; crop-damaging insect;

XX KW farm animal insect; epithelial morphogenesis; cell-matrix adhesion;

XX KW X-linked anhidrotic (hypohidrotic) ectodermal dysplasia;

XX KW X-linked anhidrotic (hypohidrotic) ectodermal dysplasia-like disorder;

XX KW sparse hair; sweat gland aberration; endotoxic shock; inflammation;

XX KW haemorrhagic necrosis of tumour; cytotoxicity; TNFv2;

XX KW obesity-linked insulin resistance; gene; ds.

XX OS Drosophila melanogaster.

XX FH Key Location/Qualifiers

XX FT CDS 634..1863

XX FT /\*tag= b

XX FT /product= "TNFv1"

XX FT /note= "Tumour necrosis factor variant 1"

XX FT misc\_difference 634..1860

XX FT /\*tag= d

XX FT sig\_peptide 634..789

XX FT /note= "Specifically claimed in claim 22"

XX FT /\*tag= a

XX FT mat\_peptide 790..1860

XX FT /\*tag= c

XX FT /label= mature\_TNFv1

XX US2002012968-A1.

XX 31-JAN-2002.

XX 20-MAR-2001; 2001US-00813329.

XX 21-MAR-2000; 2000US-0190816P.

XX (CARR/) CARROLL P M.

XX (CHEN/) CHEN J.

XX (RAMA/) RAMANATHAN C S.

XX (XIAO/) XIAO H.

PA	(GUAN/) GUAN B.	QY	541	ATCGTCATCATCATCGTCGTCTATCAACAGATCAGCATCAGCATCTCGAGGCC	600
PA	(BOWE/) BOWEN M A.	Db	541	ATCGTCATCATCATCGTCGTCTATCAACAGATCAGCATCAGCATCTCGAGGCC	600
XX					
PI	Carroll PM, Chen J, Ramanathan CS, Xiao H, Guan B, Bowen MA;	QY	601	CGGTGTCTTAAGATCCCAAGTGTTCATCAATATGACTGCCGAGACCCCTCAAGCCGTTT	660
XX	WPI; 2002-195121/25.	Db	601	CGGTGTCTTAAGATCCCAAGTGTTCATCAATATGACTGCCGAGACCCCTCAAGCCGTTT	660
XX	P-PSDB; AAU77718.				
PT	New Drosophila tumor necrosis factor molecule, useful in controlling	QY	661	ATACGCCAAGAGTGCACAGATGATGTTTTCGGGCCAAAGCAGCAGCAGCGGACC	720
PT	agriculturally important pests, e.g. comprises modifying the growth,	Db	661	ATACGCCAAGAGTGCACAGATGATGTTTTCGGGCCAAAGCAGCAGCAGCGGACC	720
PT	feeding or reproduction of crop-damaging insects or insects of farm				
XX	animals.				
PS	Claim 2; Fig 3A-C; 119pp; English.	QY	721	GCCGAGCGACGCCCGCAGCTGATCCCTCGGTTTGGGGTTTCATCGTCTGGGGCTG	780
XX		Db	721	GCCGAGCGACGCCCGCAGCTGATCCCTCGGTTTGGGGTTTCATCGTCTGGGGCTG	780
CC	The invention describes an isolated tumour necrosis factor polypeptide	QY	781	GTGTTGCCATTTCTCGACATAAGTCTGGCAGACAAACGCTGTATTCGATCTGGACAAG	840
CC	(TNF). The polypeptide and polynucleotide are useful in controlling	Db	781	GTGTTGCCATTTCTCGACATAAGTCTGGCAGACAAACGCTGTATTCGATCTGGACAAG	840
CC	agriculturally important pests, particularly by modifying the growth,				
CC	feeding and/or reproduction of crop-damaging insects or insects of farm				
CC	animals. The polypeptide and polynucleotide are useful for modulating				
CC	epithelial morphogenesis, cell-matrix adhesion in flies and mammals. Thus				
CC	the polypeptide and polynucleotide may be useful for treating,	QY	841	GAGCTCAAGAGCCTGAAGCGAGTCGTCGATAATCTCCAGCAGGTTTGGGCATAACTAT	900
CC	ameliorating or preventing X-linked anhidrotic (hypohidrotic) ectodermal	Db	841	GAGCTCAAGAGCCTGAAGCGAGTCGTCGATAATCTCCAGCAGGTTTGGGCATAACTAT	900
CC	dysplasia and X-linked anhidrotic (hypohidrotic) ectodermal dysplasia-				
CC	like disorders, e.g. sparse hair, abnormal or missing teeth or sweat	QY	901	CTGGACGAGTTTCGACGAGTTCCAAAGGAGTACGAGAATGCCCTCATCGATATCCAAAA	960
CC	gland aberrations in animals (e.g. insects and potentially humans),	Db	901	CTGGACGAGTTTCGACGAGTTCCAAAGGAGTACGAGAATGCCCTCATCGATATCCAAAA	960
CC	endotoxic shock, inflammation, haemorrhagic necrosis of tumours,				
CC	cytotoxicity and obesity-linked insulin resistance, all of which involve	QY	961	AAGGTGGATGGCTCAGCGATCAGGAGCAGCAGCAGTGGCGATGGTCTGGATTCCATT	1020
CC	TNF molecules. This sequence encodes the drosophila melanogaster tumour	Db	961	AAGGTGGATGGCTCAGCGATCAGGAGCAGCAGCAGTGGCGATGGTCTGGATTCCATT	1020
CC	necrosis factor variant 2 (TNFV2) protein, described in the invention				
XX					
SQ	Sequence 2148 BP; 645 A; 510 C; 506 G; 487 T; 0 U; 0 Other;				
		QY	1021	GCGGACGACGAGGACGACGACCTTAGCTATAGCTCTGTGGATGATGTTGGCGCAGCTAC	1080
		Db	1021	GCGGACGACGAGGACGACGACCTTAGCTATAGCTCTGTGGATGATGTTGGCGCAGCTAC	1080
		QY	1081	GAGGACTACCGGATATGTTAAATAAATCAACATCAACATCGGACCCACCCACA	1140
		Db	1081	GAGGACTACCGGATATGTTAAATAAATCAACATCAACATCGGACCCACCCACA	1140
		QY	1141	TCTGAGACCACTGCTGAGGCGGAGGCGGAGACGAGTGTCTCTGAGCTCTCAAAATGAT	1200
		Db	1141	TCTGAGACCACTGCTGAGGCGGAGGCGGAGACGAGTGTCTCTGAGCTCTCAAAATGAT	1200
		QY	1201	GACAAATGTTGATGATCTTTACAGGCTCAATCCACCAAAAAAGAGCAGGAGAGAAA	1260
		Db	1201	GACAAATGTTGATGATCTTTACAGGCTCAATCCACCAAAAAAGAGCAGGAGAGAAA	1260
		QY	1261	TCTCGCTCGATTGGCGATGTACGCAATGAGGAGCAGATATTCAGGAAATCACAAGAG	1320
		Db	1261	TCTCGCTCGATTGGCGATGTACGCAATGAGGAGCAGATATTCAGGAAATCACAAGAG	1320
		QY	1321	CTTCAGGAAAAAGTCAATCAATGAGGCAACTTTCAGGAGAGCCCTGCACCACTTCACAC	1380
		Db	1321	CTTCAGGAAAAAGTCAATCAATGAGGCAACTTTCAGGAGAGCCCTGCACCACTTCACAC	1380
		QY	1381	CGTCGCAATGCAATTCGCGCATTCGCGCATTCGCGCATTCGCGCATTCGCGCATTCGCG	1440
		Db	1381	CGTCGCAATGCAATTCGCGCATTCGCGCATTCGCGCATTCGCGCATTCGCGCATTCGCG	1440
		QY	1441	TCGAGGCGCAGCAGCCCATTTCCACTTCAGCAGCAGGCGGCTCACCAGGAAGTATGGGC	1500
		Db	1441	TCGAGGCGCAGCAGCCCATTTCCACTTCAGCAGCAGGCGGCTCACCAGGAAGTATGGGC	1500
		QY	1501	TACCATGGAGATATGATCATAGGAATGATAACGAGAAAACTCTTATCAGGACACTTTT	1560
		Db	1501	TACCATGGAGATATGATCATAGGAATGATAACGAGAAAACTCTTATCAGGACACTTTT	1560
		QY	1561	CAAAACGCGCATGGCGTCTTGGCGTGAACCAATACAGGCGCTATATACGTTATAGGCCAG	1620
		Db	1561	CAAAACGCGCATGGCGTCTTGGCGTGAACCAATACAGGCGCTATATACGTTATAGGCCAG	1620
		QY	1621	ATATGCTACAACTCGCACGACCAAGACGGATTTATCGTCTTTCAGGAGACACTCCA	1680





```
QY 421 TGTCTGTGAGGTTGTTCTGTGTGCTGTGTAGTATCTTAAATACATAGAGTGTGTTTATA 480
Db 439 TGTCTGTGAGGTTGTTCTGTGTGCTGTGTAGTATCTTAAATACATAGAGTGTGTTTATA 498
QY 481 TAAAGTCGCAAAAGCTCGAATGGAAACAGCTCTCGAGTGCCTTCGAGTGGGTGGGCAAG 540
Db 499 TAAAGTCGCAAAAGCTCGAATGGAAACAGCTCTCGAGTGCCTTCGAGTGGGTGGGCAAG 558
QY 541 ATCGTCATCATCATCATCTGTCGTCTATTAACAAGAAATCAGCATCAGCATCTGGAGGCC 600
Db 559 ATCGTCATCATCATCATCTGTCGTCTATTAACAAGAAATCAGCATCAGCATCTGGAGGCC 618
QY 601 CGGTGCTCTAAGATCCCGAGTCTCATCAATATGCACTGCCGAGACCCCTAAGCCGTTT 660
Db 619 CGGATGCTCTAAGATCCCGAGTCTCATCAATATGCACTGCCGAGACCCCTAAGCCGTTT 678
QY 661 ATAAACGCAACAGTGTCCCAACAGATGATGTTTTCGGGCCAAAGCAGCAGCGCGACC 720
Db 679 ATAAACGCAACAGTGTCCCAACAGATGATGTTTTCGGGCCAAAGCAGCAGCGCGACC 738
QY 721 GCCCAGGACGCAACCGCGAGCTGATCCCGCTGGTTTGGGGTTTCATCGGTCTGGGGCTG 780
Db 739 GCCCAGGACGCAACCGCGAGCTGATCCCGCTGGTTTGGGGTTTCATCGGTCTGGGGCTG 798
QY 781 GTGCTTGCCATTTCTGCACATAACGATCTGGCAGACAAACGCGTGTATCGCATCTGGACAAG 840
Db 799 GTGCTTGCCATTTCTGCACATAACGATCTGGCAGACAAACGCGTGTATCGCATCTGGACAAG 858
QY 841 GAGCTGMAAGAGCTGAGCGAGTCTCGATTAATCTCCAGACGCGTTTGGGCATAAACTAT 900
Db 859 GAGCTGMAAGAGCTGAGCGAGTCTCGATTAATCTCCAGACGCGTTTGGGCATAAACTAT 918
QY 901 CTGGACGAGTTCCGACGATTCGAAAGAGTACGAGATGCCCTCATCGACTATCCAAA 960
Db 919 CTGGACGAGTTCCGACGATTCGAAAGAGTACGAGATGCCCTCATCGACTATCCAAA 978
QY 961 AAGTGTGATGGCTCACCGATGAGGAGGACGACGATGGCGATGGTCTGGATTCCATT 1020
Db 979 AAGTGTGATGGCTCACCGATGAGGAGGACGACGATGGCGATGGTCTGGATTCCATT 1038
QY 1021 GCGGACGACGAGGACGACGATGATAGCTCTGTGATGATGTTGGCGGAGACTAC 1080
Db 1039 GCGGACGACGAGGACGACGATGATAGCTCTGTGATGATGTTGGCGGAGACTAC 1098
QY 1081 GAGGACTACACCGATATGTTAAATAACTCAACATGCACATACCGGCACACGCCACA 1140
Db 1099 GAGGACTACACCGATATGTTAAATAACTCAACATGCACATACCGGCACACGCCACA 1158
QY 1141 TCTGAGACCACTGTGAGGCGGAGGCGAGACGACGATGCGATCCTCAGCCTCAAAATGAT 1200
Db 1159 TCTGAGACCACTGTGAGGCGGAGGCGAGACGACGATGCGATCCTCAGCCTCAAAATGAT 1218
QY 1201 GACAATGTGTTTCGATGACTTTTACGAGTACAAATGACCAATGACCAAGAGGAGGAGAAA 1260
Db 1219 GACAATGTGTTTCGATGACTTTTACGAGTACAAATGACCAATGACCAAGAGGAGGAGAAA 1278
QY 1261 TCTCGCTCGATTGCGGATGACGAAATGAGGAGCAGAAATTAACAAGAAATCACACAGAG 1320
Db 1279 TCTCGCTCGATTGCGGATGACGAAATGAGGAGCAGAAATTAACAAGAAATCACACAGAG 1338
QY 1321 CTTTCAGGAAAAGTTCATCCATGAGGCAATCTTCCAAAGAGAGCCCTGCACACTTCCACAC 1380
Db 1339 CTTTCAGGAAAAGTTCATCCATGAGGCAATCTTCCAAAGAGAGCCCTGCACACTTCCACCTC 1398
QY 1381 CGTCGAGAAATGATTCGCCCATCGCCCATCTCTAGTCGCAAAAGCAGATCCGAGGAC 1440
Db 1399 CGTCGAGAAATGATTCGCCCATCGCCCATCTCTAGTCGCAAAAGCAGATCCGAGGAC 1458
QY 1441 TCAGAGGCGACGACCCATTTCCACTTTCAGGAGGCGCGCTCAACAAGAAAGTATGGGC 1500
Db 1459 TCAGAGGCGACGACCCATTTCCACTTTCAGGAGGCGCGCTCAACAAGAAAGTATGGGC 1518
```

```
QY 1501 TACCATGAGATATGTATAGGAAATGATAACGAGAGAACTCTTATCAGGACACTTT 1560
Db 1519 TACCATGAGATATGTATAGGAAATGATAAGGAGAGATGCTCTTATCAGGACACTTT 1578
QY 1561 CAAACGCGCATGGCGTCTTTCGACGGTGACCAATACAGGCGCTATATACGTATACGCCAG 1620
Db 1579 CAAACGCGCATGGCGTCTTTCGACGGTGACCAATACAGGCGCTATATATCGTATACGCCAG 1638
QY 1621 ATATGCTACAAACACTCGCAGCAGCAGGAGGAGGATTTATCGTCTTTCAGGAGACACTCCA 1680
Db 1639 ATATGCGGCTTACAACTCTGCACGACCAAGACGGATTTATCGTCTTTCAGGAGACACTCCA 1698
QY 1681 TTCTCTGAGTCTTGAACACGGTGCCCAACCAATCCCAATGCAATGCAATGCAATGCAATGCAAT 1740
Db 1699 TTCTCTGAGTCTTGAACACGGTGCCCAACCAATGCAATGCAATGCAATGCAATGCAATGCAAT 1758
QY 1741 ACAGTGGTCTGATCCACCTGGAAACCAACGAGAGGATCCATCTGAAGGACACTTCAACAC 1800
Db 1759 ACAGTGGTCTGATCCACCTGGAAACCAACGAGAGGATCCATCTGAAGGACACTTCAACAC 1818
QY 1801 GATCGCAATGCACTTCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1860
Db 1819 GATCGCAATGCACTTCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1878
QY 1861 TAAATGGAGAGATATATCCCGGTGAGAGATGGAATACCACTTTAAGCTTTTGTCCCGG 1920
Db 1879 TAAATGGAGAGATATATCCCGGTGAGAGATGGAATACCACTTTAAGCTTTTGTCCCGG 1938
QY 1921 CGACTGCTCGTCAATGCGATTCATCGCAGCGTGAATCCATTAAGTCTGTAGTACCTAGTC 1980
Db 1939 CGACTGCTCGTCAATGCGATTCATCGCAGCGTGAATCCATTAAGTCTGTAGTACCTAGTC 1998
QY 1981 TTAGTCACTCCAAACCTAATCTCAATCGAATCGTGCACTAGTCAAGAGACGG 2040
Db 1999 TTAGTCACTCCAAACCTAATCTCAATCGAATCGTGCACTAGTCAAGAGACGG 2058
QY 2041 AGGAAATCATATTTATTTTGTATATCTCGTTCGATCTCTAAAAGTGAATAAAAATATA 2100
Db 2059 AGGAAATCATATTTATTTTGTATATCTCGTTCGATCTCTAAAAGTGAATAAAAATATA 2118
QY 2101 TGTAGCTATTAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAACCTCGAG 2148
Db 2119 TGTAGCTATTAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAACCTCGAG 2166

RESULT 3
ABK11679
ID ABK11679 standard; DNA; 1221 Bp.
XX
AC ABK11679;
XX
DT 05-JUN-2002 (first entry)
XX
DE DNA encoding tumour necrosis factor variant 1 (TNFv1).
XX
KW Tumour necrosis factor; TNF; pesticide; crop-damaging insect;
KW farm animal insect; epithelial morphogenesis; cell-matrix adhesion;
KW X-linked anhidrotic (hypohidrotic) ectodermal dysplasia;
KW X-linked anhidrotic (hypohidrotic) ectodermal dysplasia-like disorder;
KW sparse hair; sweat gland aberration; endotoxin shock; inflammation;
KW haemorrhagic necrosis of tumour; cytotoxicity; TNFv1;
KW obesity-linked insulin resistance; gene; ds.
XX
OS Drosophila melanogaster.
XX
Key Location/Qualifiers
CDS 1..1221
FT /*tag= b
FT /product= "TNFv1"
FT /note= "tumour necrosis factor variant 1"
FT misc_difference 1..1218
FT /*tag= d
FT /note= "Specifically claimed in claim 18"
```

FT	sig_peptide	1..156	
FT	mat_peptide	/*tag= a	
FT		157..1218	
FT		/*tag= c	
FT		/label= mature_TNFv1	
XX			
PN	US2002012968-A1.		
XX			
PD	31-JAN-2002.		
XX			
PF	20-MAR-2001; 2001US-00813329.		
XX			
PR	21-MAR-2000; 2000US-0190816P.		
XX			
PA	(CARR/) CARROLL P M.		
PA	(CHEN/) CHEN J.		
PA	(RAMA/) RAMANATHAN C S.		
PA	(XIAO/) XIAO H.		
PA	(GUAN/) GUAN B.		
PA	(BOWE/) BOWEN M A.		
XX			
PI	Carroll PM, Chen J, Ramanathan CS, Xiao H, Guan B, Bowen MA;		
XX			
DR	WPI; 2002-195121/25.		
XX			
PT	New Drosophila tumor necrosis factor molecule, useful in controlling		
PT	agriculturally important pests, e.g. comprises modifying the growth,		
PT	feeding or reproduction of crop-damaging insects or insects of farm		
PT	animals.		
XX			
PS	Claim 2; Fig 2A-B; 119pp; English.		
XX			
CC	The invention describes an isolated tumor necrosis factor polypeptide		
CC	(TNF). The polypeptide and polynucleotide are useful in controlling		
CC	agriculturally important pests, particularly by modifying the growth,		
CC	feeding and/or reproduction of crop-damaging insects or insects of farm		
CC	animals. The polypeptide and polynucleotide are useful for modulating		
CC	epithelial morphogenesis, cell-matrix adhesion in flies and mammals. Thus		
CC	the polypeptide and polynucleotide may be useful for treating,		
CC	ameliorating or preventing X-linked anhidrotic (hypohidrotic) ectodermal		
CC	dysplasia and X-linked anhidrotic (hypohidrotic) ectodermal dysplasia-		
CC	like disorders, e.g. sparse hair, abnormal or missing teeth or sweat		
CC	gland aberrations in animals (e.g. insects and potentially humans),		
CC	endotoxic shock, inflammation, haemorrhagic necrosis of tumours,		
CC	cytotoxicity and obesity-linked insulin resistance, all of which involve		
CC	TNF molecules. This sequence encodes the drosophila melanogaster tumor		
CC	necrosis factor variant 1 (TNFv1) protein, described in the invention		
XX			
SQ	Sequence 1221 BP; 349 A; 323 C; 312 G; 237 T; 0 U; 0 Other;		
	Query Match	53.2%;	Score 1142.8; DB 6; Length 1221;
	Best Local Similarity	96.3%;	Pred. No. 4e-249;
	Matches 1184; Conservative	0; Mismatches	37; Indels 9; Gaps 1;
QY	634 ATGACTGCGAGACCTCAAGCCGTTTATACGCCAACAGAGTGCACACGATGATGTTT	693	
DB	1 ATGACTGCGAGACCTCAAGCCGTTTATACGCCAACAGAGTGCACACGATGATGTTT	60	
QY	694 CCGGCCAAAGCGACGACGCGCGACCCGCCAGCGACGCCGACGATCCCGCTG	753	
DB	61 CCGGCCAAAGCGACGACGCGCGACCCGCCAGCGACGCCGACGATCCCGCTG	120	
QY	754 GTTTTGGGGTTCATCGGTCTGGGGTGGTCTGTTGTCATTCGCACTAACGATCTGGCAG	813	
DB	121 GTTTTGGGGTTCATCGGTCTGGGGTGGTCTGTTGTCATTCGCACTAACGATCTGGCAG	180	
QY	814 ACAACCGGTGTATCGCATCTCGACAAGGAGCTGAAGAGCCCTGAAGCGAGTCTGCGATAAT	873	
DB	181 ACAACCGGTGTATCGCATCTCGACAAGGAGCTGAAGAGCCCTGAAGCGAGTCTGCGATAAT	240	
QY	874 CTCGACGACGCTTTGGGATAAATCTATCTGACGAGTTTCGACGAGTTCCAAAAGAGTAC	933	
DB	241 CTCGACGACGCTTTGGGATAAATCTATCTGACGAGTTTCGACGAGTTCCAAAAGAGTAC	300	

RESULT 4  
ABL21473  
ID ABL21473 standard; DNA; 978 BP.  
XX  
AC ABL21473;  
XX  
DT 26-MAR-2002 (first entry)

QY	934 GAGAAATGCCCTCATCGACTATCCAAAAAGGTGATGGCCTCACGGATGAGAGGACGAC	993
DB	301 GAGAAATGCCCTCATCGACTATCCAAAAAGGTGATGGCCTCACGGATGAGAGGACGAC	360
QY	994 GACGATGGCGATGCTCTGGATTCCATTGCGGACACGAGAGACGACGCTTAGCTATAGC	1053
DB	361 GACGATGGCGATGCTCTGGATTCCATTGCGGACACGAGAGACGACGCTTAGCTATAGC	420
QY	1054 TCTGTGGATGATGTTGGCGAGACTACGAGGACTACACCGATATGTTAAATAAATCAAC	1113
DB	421 TCTGTGGATGATGTTGGCGAGACTACGAGGACTACACCGATATGTTAAATAAATCAAC	480
QY	1114 AATGCACATACCGCACACCGCCACATCTGAGACCACTGCTGAGGGCGAGGCGAGACG	1173
DB	481 AATGCACATACCGCACACCGCCACATCTGAGACCACTGCTGAGGGCGAGGCGAGACG	540
QY	1174 GACAGTGCATCTCAGCCTCAAAATGATGACAAATGTGTTGATGATCTTTACAGCTACAA	1233
DB	541 GACAGTGCATCTCAGCCTCAAAATGATGACAAATGTGTTGATGATCTTTACAGCTACAA	600
QY	1234 GCCCACAAAAGAGAGGAGAGAAATCTCGCTCGATTGCGGATGTACGCAATGAGGAG	1293
DB	601 GCCCACAAAAGAGAGGAGAGAAATCTCGCTCGATTGCGGATGTACGCAATGAGGAG	660
QY	1294 CAGAAATTTCAAGGAAATCACACAGAGCTTCAGGAAAAGTTCATCCAATGAGGCAACTTCC	1353
DB	661 CAGAAATTTCAAGGAAATCACACAGAGCTTCAGGAAAAGTTCATCCAATGAGGCAACTTCC	720
QY	1354 AAAGAGAGCCCTGCACCACTTACACCGTGCAGAAATGCAATCCCGCATTCGCAACCTC	1413
DB	721 AAAGAGAGAAATGCAATTCCTCGCCATCGCCACCTCTAGTCCGCAAGGTGAATCTC----	775
QY	1414 CTAGTCCGCAAGCCAGATCCGAGGACTCCGAGGCCAGCAGCCCATTTCCATCTTGAGCAGC	1473
DB	776 -----TTCTTTTCAGCCAGATCCGAGGACTTCGAGGCCAGCAGCCCATTTCCATCTTGAGCAGC	831
QY	1474 AGCGCGGTTCACCAAGGAAAGTATGGGTACCATCGAGATATGTACATAGGAAATGATAAC	1533
DB	832 AGCGCGGTTCACCAAGGAAAGTATGGGTACCATCGAGATATGTACATAGGAAATGATAAC	891
QY	1534 GAGAAACCTCTTATAGGACACTTTTCAAACGCGCATGGCGTCTTGACGGTACCAAT	1593
DB	892 GAGAAACCTCTTATAGGACACTTTTCAAACGCGCATGGCGTCTTGACGGTACCAAT	951
QY	1594 ACAGGCTATATTACGTATAGCCCGACAGATGCTTACAACTCGCAGCAGCAGAACGGA	1653
DB	952 ACAGGCTATATTACGTATAGCCCGACAGATGCTTACAACTCGCAGCAGCAGAACGGA	1011
QY	1654 TTTATCGTCTTCAAGGAGACACTTCCATTCTCGAGTGTGTAACACGGTGCCCAACCAAC	1713
DB	1012 TTTATCGTCTTCAAGGAGACACTTCCATTCTCGAGTGTGTAACACGGTGCCCAACCAAC	1071
QY	1714 ATGCCACATAGGTGCACTCTGCGACACGAGTGGTCTGATTCACCTGGAAACGAAACGAG	1773
DB	1072 ATGCCACATAGGTGCACTCTGCGACACGAGTGGTCTGATTCACCTGGAAACGAAACGAG	1131
QY	1774 AGGATTCATCTGAGGACATTCACACGATGCGAATCGCAATCTGCGGGAGGGAACCAAC	1833
DB	1132 AGGATTCATCTGAGGACATTCACACGATGCGAATCGCAATCTGCGGGAGGGAACCAAC	1191
QY	1834 CGAAGCTACTTTGGCATCTTCAAGGTGTA 1863	
DB	1192 CGAAGCTACTTTGGCATCTTCAAGGTGTA 1221	

XX Drosophila melanogaster genomic polynucleotide SEQ ID NO 15892.  
DE Drosophila; developmental biology; cell signalling; insecticide;  
KW pharmaceutical; gene; ds.  
XX Drosophila melanogaster.  
OS WO200171042-A2.  
XX 27-SEP-2001.  
XX 23-MAR-2001; 2001WO-US009231.  
XX 23-MAR-2000; 2000US-0191637P.  
PR 11-JUL-2000; 2000US-00614150.  
XX (PEKE ) PE CORP NY.  
XX Venter JC, Adams M, Li PWD, Myers EW;  
XX WPI; 2001-656860/75.  
XX New isolated nucleic acid detection reagent for detecting 1000 or more  
PT genes from Drosophila and for elucidating cell signalling and cell-cell  
PT interactions.  
XX Claim 1; SEQ ID NO 15892; 21pp + Sequence Listing; English.  
XX The invention relates to an isolated nucleic acid detection reagent  
CC capable of detecting 1000 or more genes from Drosophila. The invention is  
CC useful in developmental biology and in elucidating cell signalling and  
CC cell-cell interactions in higher eukaryotes for the development of  
CC insecticides, therapeutics and pharmaceutical drugs. The invention  
CC discloses genomic DNA sequences (AB161176-AB130511), expressed DNA  
CC sequences (AB101840-AB161175) and the encoded proteins (AB57737-  
CC AB572072). The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
XX  
XX Sequence 978 BP; 289 A; 261 C; 238 G; 190 T; 0 U; 0 Other;  
XX  
XX Query Match 42.2%; Score 907.4; DB 4; Length 978;  
XX Best Local Similarity 98.0%; Pred. No. 9.5e-196;  
XX Matches 936; Conservative 0; Mismatches 1; Indels 18; Gaps 1;  
XX  
QY 927 GGAGTACGAGAAATGCCCTCATCGACTATCCAAAAGGTGGATGGCTCAGGATGAGGA 986  
DB 24 GCAGTACGAGAAATGCCCTCATCGACTATCCAAAAGGTGGATGGCTCAGGATGAGGA 83  
QY 987 GGACGACGACGATGGCGATGGTCTGGATTCCATTGGCGGACGAGGACGACGCTTAG 1046  
DB 84 GGACGACGACGATGGCGATGGTCTGGATTCCATTGGCGGACGAGGACGACGCTTAG 143  
QY 1047 CTATAGCTCTGTGATGATGTTGGCGGACGACTAGGAGCACTACCGGATGTTAAATAA 1106  
DB 144 CTATAGCTCTGTGATGATGTTGGCGGACGACTAGGAGCACTACCGGATGTTAAATAA 203  
QY 1107 ACTCAACAAATGCACATACCGGACGACCGCCACATCTGAGACCACTGCTGAGGCGGAGG 1166  
DB 204 ACTCAACAAATGCACATACCGGACGACCGCCACATCTGAGACCACTGCTGAGGCGGAGG 263  
QY 1167 CGAGACGACGATGCTATCCTCAGCCTCAATATGATGACAAATGTTTCGATGACTTTACCA 1226  
DB 264 CGAGACGACGATGCTATCCTCAGCCTCAATATGATGACAAATGTTTCGATGACTTTACCA 323  
QY 1227 CTACAATGCCCAAAAAGGAGGAGAGAAATCTCGCTCGATTCGCGATGACGAA 1286  
DB 324 CTACAATGCCCAAAAAGGAGGAGAGAAATCTCGCTCGATTCGCGATGACGAA 383  
QY 1287 TGAGGAGCAGAAATATTCAAGGAAATCACACAGAGCTTCAGGAAAGTCCAAATGAGGC 1346  
DB 384 TGAGGAGCAGAAATATTCAAGGAAATCACACAGAGCTTCAGGAAAGTCCAAATGAGGC 443

QY 1347 AACTTCCAAAGAGAGCCCTGCACCACTTCCACCGCTCGAGAATGCAATCCCGCCATCG 1406  
DB 444 AACTTCCAAAGAGAGCCCTGCACCACTTCCACCGCTCGAGAATGCAATCCCGCCATCG 503  
QY 1407 CCACCTCTAGTCCGAA-----AGCCAGATCCGAGAGACTCGAGGCC 1448  
DB 504 CCACCTCTAGTCCGAAAGGTGAATCTCTTCTTTCAGGCAGATCCGAGACTCGAGGCC 563  
QY 1449 AGCAGCCCATTTCCACTTTGAGCAGAGCGGCGCTCACCAGGAAGTATGGGCTACCATGG 1508  
DB 564 AGCAGCCCATTTCCACTTTGAGCAGAGCGGCGCTCACCAGGAAGTATGGGCTACCATGG 623  
QY 1509 AGATATGTACATAGGAATGATAACGAGAGAACTCTTATCAGGGACACTTTCAAACGCG 1568  
DB 624 AGATATGTACATAGGAATGATAACGAGAGAACTCTTATCAGGGACACTTTCAAACGCG 683  
QY 1569 CGATGGCGCTTTCAGCGTGACCAATACAGGCCTTATATCGTATACGCCAGATATGCTA 1628  
DB 684 CGATGGCGCTTTCAGCGTGACCAATACAGGCCTTATATCGTATACGCCAGATATGCTA 743  
QY 1629 CAACAACTCGCACGACGAGAACCGGATTTATCGTCTTTCAAGGAGACACTCCATTCCTGCA 1688  
DB 744 CAACAACTCGCACGACGAGAACCGGATTTATCGTCTTTCAAGGAGACACTCCATTCCTGCA 803  
QY 1689 GTGCTTGAACACGCTGCCCAACCAACATGCCACATAGGTGCAACCTGCCACACGAGTGG 1748  
DB 804 GTGCTTGAACACGCTGCCCAACCAACATGCCACATAGGTGCAACCTGCCACACGAGTGG 863  
QY 1749 TCTGATCCACCTGGAGAACGAGAACGAGAGGATCCATCTGAGGACATTCACACGATCGCAA 1808  
DB 864 TCTGATCCACCTGGAGAACGAGAACGAGAGGATCCATCTGAGGACATTCACACGATCGCAA 923  
QY 1809 TGCAGTTCGCGGAGGGGAAACACACGAGTACTTTGGCATCTTCAAGGTGTA 1863  
DB 924 TGCAGTTCGCGGAGGGGAAACACACGAGTACTTTGGCATCTTCAAGGTGTA 978  
XX  
XX RESULT 5  
XX ABL21472/c  
XX ID ABL21472 standard; DNA; 3324 BP.  
XX AC ABL21472;  
XX AC ABL21472;  
XX DT 26-MAR-2002 (first entry)  
XX XX  
DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 15889.  
XX Drosophila; developmental biology; cell signalling; insecticide;  
KW pharmaceutical; gene; ds.  
XX Drosophila melanogaster.  
OS WO200171042-A2.  
XX 27-SEP-2001.  
XX 23-MAR-2001; 2001WO-US009231.  
XX 23-MAR-2000; 2000US-0191637P.  
PR 11-JUL-2000; 2000US-00614150.  
XX (PEKE ) PE CORP NY.  
XX Venter JC, Adams M, Li PWD, Myers EW;  
XX WPI; 2001-656860/75.  
XX New isolated nucleic acid detection reagent for detecting 1000 or more  
PT genes from Drosophila and for elucidating cell signalling and cell-cell  
PT interactions.  
XX Claim 1; SEQ ID NO 15889; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent  
CC capable of detecting 1000 or more genes from *Drosophila*. The invention is  
CC useful in developmental biology and in elucidating cell signalling and  
CC cell-cell interactions in higher eukaryotes for the development of  
CC insecticides, therapeutics and pharmaceutical drugs. The invention  
CC discloses genomic DNA sequences (AB161176-AB161177), expressed DNA  
CC sequences (AB161178-AB161179) and the encoded proteins (AB161180-  
CC AB161181). The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 3324 BP; 918 A; 715 C; 681 G; 1010 T; 0 U; 0 Other;  
Query Match 36.3%; Score 779.4; DB 4; Length 3324;  
Best Local Similarity 76.4%; Pred. No. 1.7e-166;  
Matches 1184; Conservative 0; Mismatches 1; Indels 364; Gaps 4;  
QY 927 GGAGTACGAGAAATGCCCTCATCGACTATCCAAAAAGGTGGATGCCCTCAGGATGAGGA 986  
DB 2301 GCAGTACGAGAAATGCCCTCATCGACTATCCAAAAAGGTGGATGCCCTCAGGATGAGGA 2242  
QY 987 GGACGACACGATGGCGATGCTCGGATTCATTCGCGACGACGAGGACGAGCTTAG 1046  
DB 2241 GGACGACGACGATGGCGATGCTCGGATTCATTCGCGACGACGAGGACGAGCTTAG 2182  
QY 1047 CTATAGCTCTGTGGATGATGTTGGCGCAGACTACCGGAGCTACCGGATGTTAAATAA 1106  
DB 2181 CTATAGCTCTGTGGATGATGTTGGCGCAGACTACCGGAGCTACCGGATGTTAAATAA 2122  
QY 1107 ACTCAACAAATGCATACCGGACCAACCGCCACATCTGAGACCACTGCTGAGGCGAGGG 1166  
DB 2121 ACTCAACAAATGCATACCGGACCAACCGCCACATCTGAGACCACTGCTGAGGCGAGGG 2062  
QY 1167 CGAGCGCAGAGTGCATCTCAGGCTCAAAATGATGACAAATGTTGCTGATGATACGAG 1226  
DB 2061 CGAGCGCAGAGTGCATCTCAGGCTCAAAATGATGACAAATGTTGCTGATGATACGAG 2002  
QY 1227 CTACAATGCCCAAAAGAGACGAGG----- 1254  
DB 2001 CTACAATGCCCAAAAGAGACGAGGAGGTGAGCAGGATGAAATTTTCAGGGATTTA 1942  
QY 1255 -----AGAAATCTCGCTCGATGTCGCGATGATACGC 1284  
DB 1941 TCGTATTAAAGCTCTGCTTTTCTCTTCCACAGAAATCTCGCTCGATGTCGCGATGATACGC 1882  
QY 1285 AATGAGGACGAGAAATATTCAGGAAATACACAGAGCTTCAGGAAAGTCAATCAATGAG 1344  
DB 1881 AATGAGGACGAGAAATATTCAGGAAATACACAGAGCTTCAGGAAAGTCAATCAATGAG 1822  
QY 1345 GCAACTTCCAAAGAG----- 1359  
DB 1821 GCAACTTCCAAAGAGAGGTAAAGTAGACGTTCCGTTGAAACAGGCAAGAAACCAATTA 1762  
QY 1360 ----- 1359  
DB 1761 ATGAGTTTATGCTATGGTTAAAGATAGATAGTTAACTGGTAACTGAACACATATG 1702  
QY 1360 -----AGCCCTGCACCACTTCACACCGTCGCAAGATGCAATC 1397  
DB 1701 TATACATTTTCTTTTACTTGAAGCCCTGCACCACTTCACACCGTCGCAAGATGCAATC 1642  
QY 1398 CGCCATCGCACCTCTCTAGTCCGCA----- 1424  
DB 1641 CGCCATCGCACCTCTCTAGTCCGCAAGGTTGTTAGTGAAGGAAACAAATTTGGC 1582  
QY 1425 ----- 1424  
DB 1581 AAAAATCGAATACCAATCTCTATTCGCACCTCTGTTAAATTTGCAGGTGAATCTCTTCTT 1522  
QY 1425 --AGCCAGATCGAGGATCGAGGCCAGAGCCCAATTTCCATTTGAGCAGGCGCGGT 1482  
DB 1521 TCAGCCAGATCGAGGATCGAGGCCAGAGCCCAATTTCCATTTGAGCAGGCGCGGT 1462

QY 1483 CACCAAGGAAGTAGT----- 1497  
DB 1461 CACCAAGGAAGTAGTGGTCAAGGCTCAAGCGTGTACATACAACTACTAC 1402  
QY 1498 -----GGTACCATTGGAGATATGTACATAG 1522  
DB 1401 AGCAACCTAATGAAGTCACTTCTTGAATTTCTCAGGCTACCATGGAGATATGTACATAG 1342  
QY 1523 GAAATGATACGAGAGAAACTTTATCAGGAGACACTTTCAACCGCGGATGGCGCTTTGA 1582  
DB 1341 GAAATGATACGAGAGAAACTTTATCAGGAGACACTTTCAACCGCGGATGGCGCTTTGA 1282  
QY 1583 CGGTGACCAATACAGGCTATATTTACGTATATACGCCACAGATATGTACAACTCGCACG 1642  
DB 1281 CGGTGACCAATACAGGCTATATTTACGTATACGCCACAGATATGTACAACTCGCACG 1222  
QY 1643 ACAGAAAGGATTTATCGTCTTTCAAGGAGACACTCCATCTCTCGAGTGTGTTGAACAGG 1702  
DB 1221 ACCAGAAAGGATTTATCGTCTTTCAAGGAGACACTCCATCTCTCGAGTGTGTTGAACAGG 1162  
QY 1703 TGCCCAACCAATGACCAATGAGTGCACACCTGSCACACGAGTGTCTGATCCACCTGG 1762  
DB 1161 TGCCCAACCAATGACCAATGAGTGCACACCTGSCACACGAGTGTCTGATCCACCTGG 1102  
QY 1763 AACGAAACGAGAGGATCCATCTGAAAGGACATTCACAAACGATCGCAATGCGAGTTCTGCGGG 1822  
DB 1101 AACGAAACGAGAGGATCCATCTGAAAGGACATTCACAAACGATCGCAATGCGAGTTCTGCGGG 1042  
QY 1823 AGGAAACAAACCGAAGTACTTTTGGCATCTTCAAGGTGTAAATTTGGAGAGATTTATCCCG 1882  
DB 1041 AGGAAACAAACCGAAGTACTTTTGGCATCTTCAAGGTGTAAATTTGGAGAGATTTATCCCG 982  
QY 1883 GTCAGAGATGGAATACCACTTTTAAAGCTTTTGTCCCGGACCTGCTCGTGAATCGGATTC 1942  
DB 981 GTCAGAGATGGAATACCACTTTTAAAGCTTTTGTCCCGGACCTGCTCGTGAATCGGATTC 922  
QY 1943 ATGCCACGCGTGAATCCATTTAGTTCGTAGTACCTAGTCTTTAGTCACTCCAAACCTAATCT 2002  
DB 921 ATGCCACGCGTGAATCCATTTAGTTCGTAGTACCTAGTCTTTAGTCACTCCAAACCTAATCT 862  
QY 2003 CAATCGAATCGTGATATCTGATATAGTCAAGAGACGAGGAGAAATCATATTTATTTGT 2062  
DB 861 CAATCGAATCGTGATATCTGATATAGTCAAGAGACGAGGAGAAATCATATTTATTTGT 802  
QY 2063 ATATACCTCGTTCGACTCTTAAAGTGAATAAAATATATATGAGCTATTA 2111  
DB 801 ATATACCTCGTTCGACTCTTAAAGTGAATAAAATATATATGAGCTATTA 753  
RESULT 6  
AAQ21833  
ID AAQ21833 standard; DNA; 390 BP.  
XX  
AC AAQ21833;  
XX  
DT 08-JUN-1992 (first entry)  
XX  
DE Randomising oligonucleotide used in SPERT mRNA prepn.  
XX  
KW Systematic polypeptide evolution by reverse translation; SPERT;  
KW ligand binding; ss.  
XX  
OS Synthetic.  
XX  
FN W09202536-A.  
PD 20-FEB-1992.  
XX  
PF 02-AUG-1990; 90US-00561968.  
XX  
PR 02-AUG-1990; 90US-00561968.  
XX















```

FT      CDS      /product= "glycoprotein X"
FT      complement(93636..94127)
FT      /tag= 1
FT      /product= "interferon regulatory factor 4"
FT      complement(111931..112443)
FT      /tag= m
FT      /product= "capsid protein IV"
FT      complement(123808..127296)
FT      /tag= n
FT      /product= "immediate early protein"
XX
PN      W09804576-A1.
XX
PD      05-FEB-1998.
XX
PF      22-JUL-1997; 97WO-US013346.
XX
PR      25-JUL-1996; 96US-00686243.
PR      25-JUL-1996; 96US-00686349.
PR      25-JUL-1996; 96US-00686350.
PR      25-JUL-1996; 96US-00687253.
PR      05-SEP-1996; 96US-00708678.
PR      10-OCT-1996; 96US-00728323.
PR      13-NOV-1996; 96US-00747887.
PR      13-NOV-1996; 96US-00748640.
PR      29-NOV-1996; 96US-00757669.
XX
PA      (UYCO ) UNIV COLUMBIA NEW YORK.
XX
PI      Chang Y, Bohenzky RA, Russo JJ, Edelman IS, Moore PS;
XX      WPI; 1998-130615/12.
XX
PT      New nucleic acid encoding Kaposi's sarcoma associated herpes virus
PT      proteins - useful for, e.g. detecting levels of HHV8 in, and preparation
PT      of vaccines for treatment of, HIV patients.
XX
PS      Example 2; Page 135-203; 230pp; English.
XX
CC      This sequence represents the long unique region and terminal repeat of
CC      the Kaposi's sarcoma-associated herpes virus (KSHV). KSHV is also known
CC      as human herpes virus 8 (HHV8). This sequence contains the DNAs of the
CC      invention which encode KSHV polypeptides selected from: (a) viral
CC      macrophage inflammatory protein (MIP) II; (b) viral interleukin-6 (IL-6);
CC      (c) viral IRF 1; (d) complement-binding protein; glycoproteins B, M or L;
CC      (d) capsid protein IV encoded by ORF65; and (e) immediate early protein
CC      encoded by ORF73. Labelled probes for the nucleic acid, proteins encoded
CC      by it, and antibodies (Ab) specific for the proteins are useful for
CC      detecting HHV8, specifically for diagnosis of Kaposi's sarcoma, in body
CC      fluids or tissue samples. HHV8 infections can be treated with antisense
CC      or triplex forming molecules or agents that bind specifically to the
CC      protein. Ab may be used for prophylaxis or treatment of HHV8 infection,
CC      while the protein can be used in protective vaccines. Ab may also be used
CC      to differentiate between lymphomas, and HHV8 may be implicated in many
CC      other lymphoproliferative diseases such as lymphomas, leukaemia,
CC      splenomegaly and mycosis fungoides. Cells and animals containing the
CC      nucleic acid are useful for drug screening. HHV8-derived peptides can be
CC      used as targets for antiviral drugs, e.g. dihydrofolate reductase gene
CC      can be inhibited with methotrexate. These can also be used to determine
CC      the immune status of a patient infected with HIV. HHV8 derived protein
CC      viral MIP III may be used as an anti-inflammatory agent for, e.g.
CC      treating rheumatoid arthritis. This sequence is stated as containing 81
CC      open reading frames. (Updated on 27-AUG-2003 to correct OS field.)
XX
SQ      Sequence 137507 BP; 32579 A; 37795 C; 35758 G; 31375 T; 0 U; 0 Other;

```

```

Query Match      2.6%; Score 55.8; DB 2; Length 137507;
Best Local Similarity 50.6%; Pred.No. 0.12;
Matches 135; Conservative 0; Mismatches 132; Indels 0; Gaps 0;
QY      828 GCATCTGGACAGAGCTGAAGCGCTGTCGATATCTCCAGCAGCGTTT 887
DB      126310 GGATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGATCA 126251

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 14, 2004, 19:37:16 ; Search time 179 Seconds  
(without alignments)  
6659.411 Million cell update/sec

Title: US-09-813-329-5  
Perfect score: 2148  
Sequence: 1 ggcacggcgacgacgacgt.....aaaaaaaaaactcgag 2148

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents NA.\*  
1: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq.\*  
2: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq.\*  
3: /cgn2\_6/ptodata/2/ina/6A\_COMB.seq.\*  
4: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq.\*  
5: /cgn2\_6/ptodata/2/ina/PTCUS\_COMB.seq.\*  
6: /cgn2\_6/ptodata/2/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	66.8	3.1	390	US-09-197-649-7	Sequence 7, Appli
2	55.8	2.6	3489	US-08-728-323A-1	Sequence 1, Appli
3	55.8	2.6	3489	US-09-298-568-1	Sequence 1, Appli
4	55.8	2.6	3489	US-09-410-399-1	Sequence 1, Appli
5	55.8	2.6	32207	US-08-770-379-20	Sequence 20, Appli
6	55.8	2.6	32207	US-08-757-669A-20	Sequence 20, Appli
7	55.8	2.6	32207	US-09-230-371A-20	Sequence 20, Appli
8	51.6	2.4	1037	US-09-489-847-112	Sequence 112, App
9	51.4	2.4	3527	US-08-909-965C-7	Sequence 7, Appli
10	50.6	2.4	299	US-09-621-976-10211	Sequence 10211, A
11	50.6	2.4	1522	US-09-413-574-1	Sequence 1, Appli
12	50.2	2.3	1485	US-09-372-422A-39	Sequence 39, Appli
13	50.2	2.3	1736	US-09-182-816-22	Sequence 22, Appli
14	50.2	2.3	1736	US-09-182-816-24	Sequence 24, Appli
15	50.2	2.3	1736	US-09-471-528-22	Sequence 22, Appli
16	50.2	2.3	1736	US-09-471-528-24	Sequence 24, Appli
17	50.2	2.3	1736	US-09-634-530-22	Sequence 22, Appli
18	50.2	2.3	1736	US-09-634-530-24	Sequence 24, Appli
19	49.8	2.3	1133	US-09-916-204-1	Sequence 1, Appli
20	49.6	2.3	8100	US-09-554-337-4	Sequence 4, Appli
21	49.6	2.3	11517	US-07-920-281C-1	Sequence 1, Appli
22	49.6	2.3	11517	US-08-466-277-1	Sequence 1, Appli
23	49.6	2.3	15538	US-09-554-337-1	Sequence 1, Appli
24	49.4	2.3	441	US-09-601-537-10	Sequence 10, Appli
25	49.4	2.3	4121	US-09-601-537-9	Sequence 9, Appli
26	49.2	2.3	1637	US-09-205-258-178	Sequence 178, App
27	48.8	2.3	2610	US-09-545-814-1	Sequence 1, Appli

c	28	48.8	2.3	2610	4	US-09-545-814-3	Sequence 3, Appli
	29	48.6	2.3	1493	1	US-08-340-820-24	Sequence 24, Appli
	30	48.6	2.3	1493	1	US-08-593-535-24	Sequence 24, Appli
	31	48.6	2.3	2836	3	US-08-747-221B-24	Sequence 24, Appli
c	32	48.6	2.3	2836	3	US-08-747-221B-26	Sequence 26, Appli
	33	48.6	2.3	2836	3	US-09-005-051-24	Sequence 24, Appli
c	34	48.6	2.3	2836	3	US-09-005-051-26	Sequence 26, Appli
	35	48.6	2.3	2836	4	US-09-403-942F-24	Sequence 24, Appli
c	36	48.6	2.3	2836	4	US-09-403-942F-26	Sequence 26, Appli
	37	48	2.2	231	4	US-09-621-976-16317	Sequence 16317, A
	38	48	2.2	242	4	US-09-621-976-16324	Sequence 16324, A
	39	48	2.2	2329	4	US-09-800-729-11	Sequence 11, Appli
	40	47.8	2.2	746	3	US-09-013-810-1	Sequence 1, Appli
	41	47.8	2.2	1172	1	US-07-945-288-9	Sequence 9, Appli
	42	47.8	2.2	1172	1	US-08-462-831-9	Sequence 9, Appli
	43	47.8	2.2	1172	1	US-08-461-809-9	Sequence 9, Appli
	44	47.8	2.2	1172	1	US-08-461-441-9	Sequence 9, Appli
	45	47.8	2.2	1172	5	PCT-US93-08518-9	Sequence 9, Appli

ALIGNMENTS

RESULT 1  
US-09-197-649-7  
; Sequence 7, Application US/09197649  
; Patent No. 6194550  
; GENERAL INFORMATION:  
; APPLICANT: Gold, Larry  
; APPLICANT: Tuerk, Craig  
; APPLICANT: Pribnow, David  
; APPLICANT: Smith, Jonathan D.  
; TITLE OF INVENTION: Systematic Polypeptide Evolution by Reverse Translation  
; FILE REFERENCE: NEX02/CJ-CON  
; CURRENT APPLICATION NUMBER: US/09/197,649  
; CURRENT FILING DATE: 1998-11-23  
; EARLIER APPLICATION NUMBER: 07/829,461  
; EARLIER FILING DATE: 1992-01-31  
; EARLIER APPLICATION NUMBER: 07/739,055  
; EARLIER FILING DATE: 1991-08-01  
; EARLIER APPLICATION NUMBER: 07/561,968  
; EARLIER FILING DATE: 1990-08-02  
; NUMBER OF SEQ ID NOS: 26  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 7  
; LENGTH: 390  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Sequence  
; OTHER INFORMATION: having a 120 repeat of ACG flanked by fixed  
; OTHER INFORMATION: fragments having NcoI restriction sites.  
US-09-197-649-7

Query Match	3.1%	Score 66.8	DB 3	Length 390
Best Local Similarity	48.9%	Pred. No. 3.9e-07		
Matches 179	Conservative 0	Mismatches 187	Indels 0	Gaps 0
Qy	888	GGGCTAAATCTCTGGACGAGTTCGACGAGTTCCAAAAGGAGTAGTACGAGAATGCCCTCAT	947	
Db	2	GGCCATGGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG	61	
Qy	948	CGACTATCCAAAAAGTGGATGGCCTCACGGATGAGGAGGACGACGACGATGGCGATGG	1007	
Db	62	CGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG	121	
Qy	1008	TCTGGATTCCATTGCGGACGACGACGACGACGACGACGACGACGACGACGACGACGACG	1067	
Db	122	CGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG	181	
Qy	1068	TGGCGGACGACTACGAGGACTACACCGATATGTTAAATAAATCAACAATGCATACCGG	1127	
Db	182	CGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG	241	









```
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FITZPATRICK, CELLA, HARPER AND SCINTO
; STREET: 277 Park Avenue
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10172-0194
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/909,965C
; FILING DATE: August 12, 1997
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 322745/95
; APPLICATION NUMBER: PCT/JP96/03630
; FILING DATE: 12-No. 5936078-1995
; FILING DATE: 12-Dec-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Lawrence S. Perry
; REGISTRATION NUMBER: 31865
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-2400
; TELEFAX: 212-758-2982
; TELEX: 236262
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3527 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; ORIGINAL SOURCE:
; ORGANISM: human
; IMMEDIATE SOURCE:
; CLONE: F998
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 140 to 1084
; IDENTIFICATION METHOD: by experiment
US-08-909-965C-7

Query Match 2.4%; Score 51.4; DB 2; Length 3527;
Best Local Similarity 69.3%; Pred. No. 0.0078;
Matches 70; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

Qy 2042 GGAAATCATATTTATTTTGTATATCTCGTTTCGACTCTAAAGTGAATAAAATATAT 2101
Db 3419 GTAATAAAATCAATATGTATATAATCTCGTGAATCTAAAGTGAATAAAATATAT 3478

Qy 2102 GTAGCTATTAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2142
Db 3479 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3519

RESULT 10
US-09-621-976-10211
; Sequence 10211, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
```

```
; SEQ ID NO 10211
; LENGTH: 299
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-10211

Query Match 2.4%; Score 50.6; DB 4; Length 299;
Best Local Similarity 65.5%; Pred. No. 0.0036;
Matches 74; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

Qy 2030 TCAGAACGAGGAGAAATCATATTTTGTATATCTCGTTTCGACTCTAAAGTGA 2089
Db 179 TTAAGCAGTATGGGAACCAATTTTGTAAAGTAATATTTTCAACCTTTAAAAAAA 238

Qy 2090 ATAAAAATATATGTAGCTATTAAAAAATAAAAAAAAAAAAAAAAAAAAAA 2142
Db 239 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 291

RESULT 11
US-09-413-574-1
; Sequence 1, Application US/09413574
; Patent No. 6235972
; GENERAL INFORMATION:
; APPLICANT: Mahajan, Pramod B.
; APPLICANT: Tagliani, Laura
; TITLE OF INVENTION: Maize Rad23 Genes and Uses Thereof
; FILE REFERENCE: 0964
; CURRENT APPLICATION NUMBER: US/09/413,574
; CURRENT FILING DATE: 1999-10-06
; EARLIER APPLICATION NUMBER: 60/109,728
; EARLIER FILING DATE: 1998-11-23
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 1522
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (58) ... (1272)
US-09-413-574-1

Query Match 2.4%; Score 50.6; DB 3; Length 1522;
Best Local Similarity 57.1%; Pred. No. 0.0081;
Matches 92; Conservative 0; Mismatches 69; Indels 0; Gaps 0;

Qy 1982 TAGTCATCCAAACCTAATCTCAATCGGAATCGTGCATATGCAATGCTCAGAACGGA 2041
Db 1360 TCGTCTATGCTGCTGCTGACTATCTTTTATTTCCATATATATTTGTCGAATGCTTT 1419

Qy 2042 GGAATAATCATATTTATTTTGTATATCTCGTTTCGACTCTAAAGTGAATAAAATATAT 2101
Db 1420 CTAAGTCAATATTAATTCATATCAACGCTGTACACCGTGTAAATAAAAAAAAAAAAA 1479

Qy 2102 GTAGCTATTAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2142
Db 1480 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1520

RESULT 12
US-09-372-422A-39
; Sequence 39, Application US/09372422A
; Patent No. 6313375
; GENERAL INFORMATION:
; APPLICANT: Rudolf Jung
; APPLICANT: Francois Barrieu
; TITLE OF INVENTION: Maize Aquaporins and Uses Thereof
; FILE REFERENCE: 0919
; CURRENT APPLICATION NUMBER: US/09/372,422A
; CURRENT FILING DATE: 1999-08-11
; PRIOR APPLICATION NUMBER: US 60/098,692
; PRIOR FILING DATE: 1998-08-31
```

```
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: Fast-SEQ for Windows Version 3.0
; SEQ ID NO 39
; LENGTH: 1485
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (133)...(1015)
US-09-372-422A-39

Query Match          2.3%; Score 50.2; DB 4; Length 1485;
Best Local Similarity 63.9%; Pred. No. 0.01;
Matches 76; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

Qy 2022 TGCATTAGTCAGAGACGGAGGAATATCATATTTTGTATATCTCGTTCGACTCTA 2081
Db 1358 TGCATGTGTGGAATCGCGAACACAAATATATATATGCGTTTCTTAAAAAAA 1417

Qy 2082 AAAAGTGAATAAAATATATGTAGCTATTAAAAAATATATATATATATATATAT 2140
Db 1418 AAAAATAAAATATATATATATATATATATATATATATATATATATATATAT 1476

RESULT 13
US-09-182-816-22
; Sequence 22, Application US/09182816
; Patent No. 6143542
; GENERAL INFORMATION:
; APPLICANT: Wisniewski, Nancy
; APPLICANT: Silver, Gary M.
; APPLICANT: Lo, Katherine C.
; APPLICANT: Brandt, Kevin S.
; TITLE OF INVENTION: NOVEL FLEA EPOXIDE HYDROLASE NUCLEIC ACID MOLECULES,
; FILE REFERENCE: FC-3-C1
; CURRENT APPLICATION NUMBER: US/09/182,816
; CURRENT FILING DATE: 1998-10-29
; EARLIER APPLICATION NUMBER: 08/989,510
; EARLIER FILING DATE: 1997-12-12
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 22
; LENGTH: 1736
; TYPE: DNA
; ORGANISM: Ctenocephalides felis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (159)..(1553)
US-09-182-816-22

Query Match          2.3%; Score 50.2; DB 3; Length 1736;
Best Local Similarity 68.0%; Pred. No. 0.011;
Matches 70; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

Qy 2040 GAGGAAATCATATTTTGTATATATCTCGTTCGACTCTAAAAAGTGAATAAATAT 2099
Db 1617 GAAATAAACGATATGGATTTTATTTCAAACTTGTCAATATATATATATATATAT 1676

Qy 2100 ATGTAGCTATTAAAAAATATATATATATATATATATATATATATATATATAT 2142
Db 1677 AAAAAAATATATATATATATATATATATATATATATATATATATATATATAT 1719

RESULT 14
US-09-182-816-24/c
; Sequence 24, Application US/09182816
; Patent No. 6143542
; GENERAL INFORMATION:
; APPLICANT: Wisniewski, Nancy
; APPLICANT: Silver, Gary M.
; APPLICANT: Lo, Katherine C.
; APPLICANT: Brandt, Kevin S.
; TITLE OF INVENTION: NOVEL FLEA EPOXIDE HYDROLASE NUCLEIC ACID MOLECULES,
; FILE REFERENCE: FC-3-C1
; CURRENT APPLICATION NUMBER: US/09/182,816
; CURRENT FILING DATE: 1998-10-29
; EARLIER APPLICATION NUMBER: 08/989,510
; EARLIER FILING DATE: 1997-12-12
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 24
; LENGTH: 1736
; TYPE: DNA
; ORGANISM: Ctenocephalides felis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (159)..(1553)
US-09-182-816-24

Query Match          2.3%; Score 50.2; DB 3; Length 1736;
Best Local Similarity 68.0%; Pred. No. 0.011;
Matches 70; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

Qy 2040 GAGGAAATCATATTTTGTATATATCTCGTTCGACTCTAAAAAGTGAATAAATAT 2099
Db 1617 GAAATAAACGATATGGATTTTATTTCAAACTTGTCAATATATATATATATATAT 1676

Qy 2100 ATGTAGCTATTAAAAAATATATATATATATATATATATATATATATATATAT 2142
Db 1677 AAAAAAATATATATATATATATATATATATATATATATATATATATATATAT 1719

RESULT 15
US-09-471-528-22
; Sequence 22, Application US/09471528
; Patent No. 6153397
; GENERAL INFORMATION:
; APPLICANT: Wisniewski, Nancy
; APPLICANT: Silver, Gary M.
; APPLICANT: Lo, Katherine C.
; APPLICANT: Brandt, Kevin S.
; TITLE OF INVENTION: FLEA EPOXIDE HYDROLASE PROTEINS AND USSES THEREOF
; FILE REFERENCE: FC-3-C1-1
; CURRENT APPLICATION NUMBER: US/09/471,528
; CURRENT FILING DATE: 1999-12-27
; EARLIER APPLICATION NUMBER: 09/182,816
; EARLIER FILING DATE: 1998-10-29
; EARLIER APPLICATION NUMBER: 08/989,510
; EARLIER FILING DATE: 1997-12-12
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 22
; LENGTH: 1736
; TYPE: DNA
; ORGANISM: Ctenocephalides felis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (159)..(1553)
US-09-471-528-22

Query Match          2.3%; Score 50.2; DB 3; Length 1736;
Best Local Similarity 68.0%; Pred. No. 0.011;
Matches 70; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

Qy 2040 GAGGAAATCATATTTTGTATATATCTCGTTCGACTCTAAAAAGTGAATAAATAT 2099
Db 1617 GAAATAAACGATATGGATTTTATTTCAAACTTGTCAATATATATATATATATAT 1676

Qy 2100 ATGTAGCTATTAAAAAATATATATATATATATATATATATATATATATATAT 2142
Db 1677 AAAAAAATATATATATATATATATATATATATATATATATATATATATATAT 1719

Search completed: September 15, 2004, 02:20:54
Job time : 181 secs
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 15, 2004, 00:00:21 ; Search time 4276 Seconds  
(without alignments)  
2527.528 Million cell updates/sec

Title: US-09-813-329-5  
Perfect score: 2148  
Sequence: 1 ggcacggcggaacggacgt.....aaaaaaaaaactcgag 2148

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3304383 seqs, 2515761380 residues

Total number of hits satisfying chosen parameters: 6608766

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database : Published Applications NA:  
1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq.\*  
3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq.\*  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq.\*  
6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq.\*  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq.\*  
10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq.\*  
11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq.\*  
12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq.\*  
13: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq.\*  
14: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq.\*  
15: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq.\*  
16: /cgn2\_6/ptodata/2/pubpna/US10C\_PUBCOMB.seq.\*  
17: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq.\*  
18: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq.\*  
19: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2148	100.0	2148	9	US-09-813-329-5
2	2112.8	98.4	2166	9	US-09-813-329-1
3	1142.8	53.2	1221	9	US-09-813-329-3
4	66.8	3.1	390	9	US-09-790-399-7
5	57.2	2.7	966	13	US-10-425-114-7734
6	57.2	2.7	1269	13	US-10-424-599-46034
7	55.8	2.6	3489	13	US-09-894-273-1
8	55.8	2.6	3489	15	US-10-294-804-1
c 9	54.4	2.5	5306	15	US-10-311-455-483
10	53.6	2.5	1915	15	US-10-225-567A-606
11	53.2	2.5	2190	13	US-10-424-599-9903
12	52.8	2.5	2190	13	US-10-424-599-82347
c 13	52.8	2.5	372	13	US-10-424-599-38982
14	52.6	2.4	1898	16	US-10-264-049-287

c 15	52.4	2.4	351	10	US-09-814-353-17579	Sequence 17579, A
16	52.4	2.4	479	10	US-09-918-995-9339	Sequence 9339, Ap
17	52.4	2.4	553	13	US-10-424-599-6106	Sequence 6106, Ap
18	52.4	2.4	1506	17	US-10-437-963-40658	Sequence 40658, A
c 19	52.4	2.4	5858	15	US-10-311-455-1285	Sequence 1285, Ap
20	52	2.4	399	13	US-10-424-599-26900	Sequence 26900, A
21	51.8	2.4	3530	17	US-10-437-963-24712	Sequence 24712, A
c 22	51.6	2.4	279	13	US-10-424-599-93967	Sequence 93967, A
23	51.6	2.4	417	13	US-10-424-599-116955	Sequence 116955, A
c 24	51.6	2.4	550	10	US-09-918-995-31149	Sequence 31149, A
c 25	51.6	2.4	603	17	US-10-021-323-5966	Sequence 5966, Ap
c 26	51.6	2.4	672	15	US-10-198-846-8818	Sequence 8818, Ap
c 27	51.6	2.4	1029	17	US-10-767-701-11693	Sequence 11693, A
28	51.6	2.4	1037	13	US-10-351-334-112	Sequence 112, App
c 29	51.6	2.4	1309	15	US-10-106-698-2117	Sequence 2117, Ap
c 30	51.4	2.4	337	9	US-09-960-352-6975	Sequence 6976, Ap
31	51.4	2.4	1333	13	US-10-424-599-59340	Sequence 59340, A
32	51.4	2.4	1933	13	US-10-424-599-33751	Sequence 33751, A
33	51.2	2.4	302	13	US-10-424-599-17202	Sequence 17202, A
c 34	51.2	2.4	2122	16	US-10-374-780A-884	Sequence 884, App
c 35	51.2	2.4	14006	15	US-10-311-455-1931	Sequence 1931, Ap
c 36	51	2.4	375	9	US-09-960-352-15014	Sequence 15014, A
c 37	51	2.4	418	13	US-10-424-599-118975	Sequence 118975, A
c 38	51	2.4	472	17	US-10-437-963-57216	Sequence 57216, A
c 39	51	2.4	493	13	US-10-424-599-95204	Sequence 95204, A
c 40	51	2.4	517	17	US-10-021-323-11054	Sequence 11054, A
41	51	2.4	527	9	US-09-864-761-16481	Sequence 16481, A
c 42	51	2.4	622	13	US-10-424-599-3380	Sequence 3380, Ap
c 43	51	2.4	3510	10	US-09-766-511B-31	Sequence 31, Appl
44	51	2.4	3510	15	US-10-241-220-38	Sequence 38, Appl
45	51	2.4	3552	10	US-09-814-353-19891	Sequence 19891, A

ALIGNMENTS

RESULT 1

US-09-813-329-5  
; Sequence 5, Application US/09813329  
; Patent No. US20020012968A1  
; GENERAL INFORMATION:  
; APPLICANT: Bristol-Myers Subb Company  
; TITLE OF INVENTION: No. US20020012968A1el Drosophila Tumor Necrosis Factor Class Mole  
; TITLE OF INVENTION: Variants Thereof  
; FILE REFERENCE: D0016.np  
; CURRENT APPLICATION NUMBER: US/09/813,329  
; CURRENT FILING DATE: 2001-03-20  
; PRIOR APPLICATION NUMBER: 60/190,816  
; PRIOR FILING DATE: 2000-03-21  
; NUMBER OF SEQ ID NOS: 65  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 5  
; LENGTH: 2148  
; TYPE: DNA  
; ORGANISM: Drosophila melanogaster  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (634)..(1860)  
US-09-813-329-5

Query Match	100.0%	Score 2148;	DB 9;	Length 2148;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 2148;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	GGCAGCGGCGACGGACGCTTTAAAGTGAGAAAAAAGAACCGGTAAATCAGAGATCCCAAG	60	
Db	1	GGCAGCGGCGGCGACGGACGCTTTAAAGTGAGAAAAAAGAACCGGTAAATCAGAGATCCCAAG	60	
Qy	61	CAAGCGCGTGGCGATGATAGCGAGAAAAAGCTATCCGTTTCAAGTAACTACTTAC	120	
Db	61	CAAGCGCGTGGCGATGATAGCGAGAAAAAGCTATCCGTTTCAAGTAACTACTTAC	120	
Qy	121	CAAGATTGAATTTCCGCATCGGCAATTAATAAATAACATAAGTCAACTCGTCCACT	180	



; CURRENT FILING DATE: 2001-03-20									
; PRIOR APPLICATION NUMBER: 60/190,816									
; PRIOR FILING DATE: 2000-03-21									
; NUMBER OF SEQ ID NOS: 65									
; SOFTWARE: PatentIn version 3.0									
; SEQ ID NO 1									
; LENGTH: 2166									
; TYPE: DNA									
; ORGANISM: Drosophila melanogaster									
; FEATURE:									
; NAME/KEY: CDS									
; LOCATION: (652)..(1878)									
US-09-813-329-1									
Query Match									
Best Local Similarity 98.4%; Score 2112.8; DB 9; Length 2166;									
Matches 2126; Conservative 0; Mismatches 22; Indels 0; Gaps 0;									
Qy	1	GGCAGCGCGAAACGGACGTTTAAAGTGAGAAAAAGAAACCGGTAAATCAGAGATCCCAAG	60						
Db	19	GGCAGCGCGAAACGGACGTTTAAAGTGAGAAAAAGAAACCGGTAAATCAGAGATCCCAAG	78						
Qy	61	CAAGCGCTGCGTGATATAGCGAAGAAAAAGCTATCCGTTTCAGTTAACTACTTAC	120						
Db	79	CAAGCGCTGCGTGATATAGCGAAGAAAAAGCTATCCGTTTTCAGTTAACTACTTAC	138						
Qy	121	CAAGATTGAATTTCCGCATCGGCAAAATTAATAAATAACATAAGTGCAACTCGTCCACT	180						
Db	139	CAAGATTGAATTTCCGCATCGGCAAAATTAATAAATAACATAAGTGCAACTCGTCCACT	198						
Qy	181	GTGTGTTGTTTTTTTTTTTTTTTTTTTTTTCGCTGCTGCTTTATCGCAAAACAAGAAC	240						
Db	199	GTGTGTTGTTTTTTTTTTTTTTTTTTTTTTCGCTGCTGCTTTATCGCAAAACAAGAAC	258						
Qy	241	TGATAAACTAGAAAATATCTTGAAAACTTTGTTTCGCGCTTTTCTTTTCTTAATTGCC	300						
Db	259	TGATAAACTAGAAAATATCTTGAAAACTTTGTTTTCGCGCTTTTCTTTTCTTAATTGCC	318						
Qy	301	GATCGGAAGAGAAAAACAAGCAGTACAAAAACAAGTGTGTAATAACAATCTGAAAG	360						
Db	319	GATCGGAAGAGAAAAACAAGCAGTACAAAAACAAGTGTGTAATAACAATCTGAAAG	378						
Qy	361	GGCACCATCAGCAGCCGAGGGTTTATCTATATAGATGTCGCGAGCTTATCATCTCATGC	420						
Db	379	GGCACCATCAGCAGCCGAGGGTTTATCTATATAGATGTCGCGAGCTTATCATCTCATGC	438						
Qy	421	TGCTGTGAGGTTGTTCTGTGTGCTCGTGAGTATCTTAAATACATAGAGTGTGTTCTATA	480						
Db	439	TGCTGTGAGGTTGTTCTGTGTGCTCGTGAGTATCTTAAATACATAGAGTGTGTTCTATA	498						
Qy	481	TAAAGTGGCAAAAGCTCGATTGGAACAAGCTGTGAGTGCCCTTGAGTGGGTGGCAAG	540						
Db	499	TAAAGTGGCAAAAGCTCGATTGGAACAAGCTGTGAGTGCCCTTGAGTGGGTGGCAAG	558						
Qy	541	ATCGTCATCATCATCGTCGTATTATCAACAGAATCAGCATCAGCATCTGGAGGCC	600						
Db	559	ATCGTCATCATCATCGTCGTATTATCAACAGAATCAGCATCAGCATCTGGAGGCC	618						
Qy	601	CGGTTGCTCTAAGATCCCAGTGTTTCATCAATTATGACTGCCGAGACCCCTCAAGCCGTTT	660						
Db	619	CGGATGCTCTAAGATCCCAGTGTTTCATCAATTATGACTGCCGAGACCCCTCAAGCCGTTT	678						
Qy	661	ATAAGCCAAACGAGTGCCAAACGATGATGTTTTCCGGCCAAAGCAGCAGCAGCGCAGC	720						
Db	679	ATAAGCCAAACGAGTGCCAAACGATGATGTTTTCCGGCCAAAGCAGCAGCAGCGCAGC	738						
Qy	721	GCCCAGCAGCACCCCGCAGCTGATCCCTCGTGTGTTTTGGGGTTTCATCGGCTCGGGGCTG	780						
Db	739	GCCCAGCAGCACCCCGCAGCTGATCCCTCGTGTGTTTTGGGGTTTCATCGGCTCGGGGCTG	798						
Qy	781	GTGCTTGCCATTCTCGCACTAACGATCTGGCAGACAAACGCGTGATCGCATCTGGACAAAG	840						
Db	799	GTGCTTGCCATTCTCGCACTAACGATCTGGCAGACAAACGCGTGATCGCATCTGGACAAAG	858						

Qy	841	GAGCTGAAGAGCCCTGAAGCGAGTCGTGATAAATCTCCAGCAGCGTTTTGGGCATAAACTAT	900
Db	859	GAGCTGAAGAGCCCTGAAGCGAGTCGTGCGATAAATCTCCAGCAGCGTTTTGGGCATAAACTAT	918
Qy	901	CTGGACGAGTTTCGACGAGTTTCCAAAAGAGGTACAGAAATGCCCTCATCGACTATCCAAAA	960
Db	919	CTGGACGAGTTTCGACGAGTTTCCAAAAGAGGTACAGAAATGCCCTCATCGACTATCCAAAA	978
Qy	961	AAGGTGATGCGCTCACGGATGAGGAGCAGCAGCATGGCATGTGCTGGATTCATTT	1020
Db	979	AAGGTGATGCGCTCACGGATGAGGAGCAGCAGCATGGCATGTGCTGGATTCATTT	1038
Qy	1021	GCGACGACGAGGACGACGAGCTTTAGCTATAGCTCTCTGGATGATGTTGGCGCAGACTAC	1080
Db	1039	GCGACGACGAGGACGACGAGCTTTAGCTATAGCTCTCTGGATGATGTTGGCGCAGACTAC	1098
Qy	1081	GAGGACTACACCCGATATGTTAAATAAACTCAAATGCACATACCCGACACACGCCCA	1140
Db	1099	GAGGACTACACCCGATATGTTAAATAAACTCAAATGCACATACCCGACACACGCCCA	1158
Qy	1141	TCTGAGACCACTGCTGAGGCGAGCGAGCAGTGCATCCTCAGCCTCAAAATGAT	1200
Db	1159	TCTGAGACCACTGCTGAGGCGAGCGAGCAGTGCATCCTCAGCCTCAAAATGAT	1218
Qy	1201	GACAAATGTGTTTCCGATGATTTTACCAGCTACAATGCCCAAAAAAGAGGAGAGAAAA	1260
Db	1219	GACAAATGTGTTTCCGATGATTTTACCAGCTACAATGCCCTCAAAAAAGAGGAGAGAAAA	1278
Qy	1261	TCTCGCTCGATTTGCCGATGTACGCAATGAGGAGCAGAAATATTTCAAGGAAATTCACACAGAG	1320
Db	1279	TCTCGCTCGATTTGCCGATGTACGCAATGAGGAGCAGAAATATTTCAAGGAAATTCACACAGAG	1338
Qy	1321	CTTCAGGAAAAGTCCAAATGAGGCAACTTCCAAAGAGAGCCCTGCGACACTTTCACCTC	1380
Db	1339	CTTCAGGAAAAGTCCAAATGAGGCAACTTCCAAAGAGAGCCCTGCGACACTTTCACCTC	1398
Qy	1381	CGTCGCAAGATGCAATTTCCCGCCATCGCCACCTCTAGTCGCGAAAGCCAGATCCGAGGAC	1440
Db	1399	CGTCGCAAGATGCAATTTCCCGCCATCGCCACCTCTAGTCGCGAAAGCCAGATCCGAGGAC	1458
Qy	1441	TCGAGGCGCAGCAGCCCATTTCCATTGAGCAGCAGCGCGGTTCACCAAGGAAGTATGGGC	1500
Db	1459	TCGAGGCGCAGCAGCCCATTTCCATTGAGCAGCAGCGCGGTTCACCAAGGAAGTATGGGC	1518
Qy	1501	TACCATCGAGATGTACATAGGAAATGATAACAGAGAGAACTCTTATCAGGGACACTTT	1560
Db	1519	TACCATCGAGATGTACATAGGAAATGATAACAGAGAGAACTCTTATCAGGGACACTTT	1578
Qy	1561	CAAAACGCGCGATGCGCTTTTGAACGTCACCAATACAGGCCCTATATTACGTATACGCCAG	1620
Db	1579	CAAAACGCGCGATGCGCTTTTGAACGTCACCAATACAGGCCCTATATTACGTATACGCCAG	1638
Qy	1621	ATATGCTACAACTACGACGACGAGATTTATTCGTTCTTTCAAGGAGACACTTCCA	1680
Db	1639	ATATGCGGCTACAACTACGACGACGAGATTTATTCGTTCTTTCAAGGAGACACTTCCA	1698
Qy	1681	TTCCCTGAGTGTGTTGAACGCGTGCACCAACCAATGCCACATAGGTGCAACCTGCCAC	1740
Db	1699	TTCCCTGAGTGTGTTGAACGCGTGCACCAACCAATGCCACATAGGTGCAACCTGCCAC	1758
Qy	1741	ACGAGTGTCTGATTCACCTGGAAACGAAACGAGAGATCCATCTGAAAGGACATTTCACAAC	1800
Db	1759	ACGAGTGTCTGATTCACCTGGAAACGAAACGAGAGATCCATCTGAAAGGACATTTCACAAC	1818
Qy	1801	GATCGCAATGAGTTCTGCGGAGGGGAAACAAACGAAAGCTACTTTTGGGATCTTCAAGGTG	1860
Db	1819	GATCGCAATGAGTTCTGCGGAGGGGAAACAAACGAAAGCTACTTTTGGGATCTTCAAGGTG	1878
Qy	1861	TAAATTCGAGAGATTTCCCGGTCAGAAAGTGAATACAGTTTAAAGCTTTTCTGCCCG	1920
Db	1879	TAAATTCGAGAGATTTCCCGGTCAGAAAGTGAATACAGTTTAAAGCTTTTCTGCCCG	1938



QY 1921 CGACTGCTCGTGAATCGGATTCATCGCAGCGTGAATCCATTAGTTCGTAGTACCTAGTC 1980  
Db 1939 CGACTGCTCGTGAATCGGATTCATCGCAGCGTGAATCCATTAGTTCGTAGTACCTAGTC 1998  
QY 1981 TTAGTCACCTCCAAACCTAATCTCAATCGGAATCGTGCACTACTGCAATAGTCAGAGACGG 2040  
Db 1999 TTAGTCACCTCCAAACCTAATCTCAATCGGAATCGTGCACTACTGCAATAGTCAGAGACGG 2058  
QY 2041 AGGAAATCATATTTATTTTGTATATCTGCTGCACTCTAAAAGTGAATAAATATA 2100  
Db 2059 AGGAAATCATATTTATTTTGTATATCTGCTGCACTCTAAAAGTGAATAAATATA 2118  
QY 2101 TGTAGCTATTAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAACTCGAG 2148  
Db 2119 TGTAGCTATTAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAACTCGAG 2166

RESULT 3  
US-09-813-329-3  
; Sequence 3, Application US/09813329  
; Patent No. US20020012968A1  
; GENERAL INFORMATION:  
; APPLICANT: Bristol-Myers Squibb Company  
; TITLE OF INVENTION: No. US20020012968A1el Drosophila Tumor Necrosis Factor Class Mole  
; FILE REFERENCE: Variants Thereof  
; CURRENT APPLICATION NUMBER: US/09/813,329  
; CURRENT FILING DATE: 2001-03-20  
; PRIOR APPLICATION NUMBER: 60/190,816  
; PRIOR FILING DATE: 2000-03-21  
; NUMBER OF SEQ ID NOS: 65  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 3  
; LENGTH: 1221  
; TYPE: DNA  
; ORGANISM: Drosophila melanogaster  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(1218)  
US-09-813-329-3

Query Match 53.2%; Score 1142.8; DB 9; Length 1221;  
Best Local Similarity 96.3%; Pred. No. 4.4e-298;  
Matches 1184; Conservative 0; Mismatches 37; Indels 9; Gaps 1;

QY 634 ATGACTGCCGAGACCTCAAGCCGTTTATAACGCCAACGAGTGCACAGATGATGTTTT 693  
Db 1 ATGACTGCCGAGACCTCAAGCCGTTTATAACGCCAACGAGTGCACAGATGATGTTTT 60  
QY 694 CCGGCCAAAGCGACCGAGCCGCCAGCGACCGCCAGCGACCGCCAGCTGATCCCCCTG 753  
Db 61 CCGGCCAAAGCGACCGAGCCGCCAGCGACCGCCAGCGACCGCCAGCTGATCCCCCTG 120  
QY 754 GTTTTGGGTTTCATCGCTCGTGGGCTGGTGGTGGCTTCTCGCACTAACGATCTGGCAG 813  
Db 121 GTTTTGGGTTTCATCGCTCGTGGGCTGGTGGTGGCTTCTCGCACTAACGATCTGGCAG 180  
QY 814 ACAACGGCTGTATCGCATCTGGACAAGAGCTGAAGAGCTGAAGCGAGTCTCGATAAT 873  
Db 181 ACAACGGCTGTATCGCATCTGGACAAGAGCTGAAGAGCTGAAGCGAGTCTCGATAAT 240  
QY 874 CTCAGCAGGCTTTGGGCATAAATACTATCTGGAGAGTTGCAAGTTCCTCAAAAGGAGTAC 933  
Db 241 CTCAGCAGGCTTTGGGCATAAATACTATCTGGAGAGTTTCCAGGAGTTCCTCAAAAGGAGTAC 300  
QY 934 GAGATCCCTCATCGACTATCAAAAAAGTGGTGGCTTCCAGTACGATGAGGAGCGAC 993  
Db 301 GAGATCCCTCATCGACTATCAAAAAAGTGGTGGCTTCCAGTACGATGAGGAGCGAC 360  
QY 994 GAGATGGCGATGCTCGGATTCATTTGCGGACGAGGACGACGACGCTTAGCTATAGC 1053  
Db 361 GAGATGGCGATGCTCGGATTCATTTGCGGACGAGGACGACGACGCTTAGCTATAGC 420

QY 1054 TCTGTGATGATGTTGGCGGAGACTACGAGGACTACACCGATATGTTAAATAAATCAAC 1113  
Db 421 TCTGTGATGATGTTGGCGGAGACTACGAGGACTACACCGATATGTTAAATAAATCAAC 480  
QY 1114 AATGCACATACCGCCACCGCCACATCTGAGACCACTCTGCTGAGGGCGAGGAGACG 1173  
Db 481 AATGCACATACCGCCACCGCCACATCTGAGACCACTCTGCTGAGGGCGAGGAGACG 540  
QY 1174 GACATGTCATCTCAGCCTCAAAATGATGACAAATGTGTTCGATGACTTTTACCAGTACAAT 1233  
Db 541 GACATGTCATCTCAGCCTCAAAATGATGACAAATGTGTTCGATGACTTTTACCAGTACAAT 600  
QY 1234 GCCCACAATAAGAGCAGGAGAGAAATCTCGCTCGATTCGCCGATGTACGCAATGAGGAG 1293  
Db 601 GCCCACAATAAGAGCAGGAGAGAAATCTCGCTCGATTCGCCGATGTACGCAATGAGGAG 660  
QY 1294 CAGAATATTCAGGAAATCACACAGAGCTTCAGGAAAGTTCATCCATGAGGCAACTTCC 1353  
Db 661 CAGAATATTCAGGAAATCACACAGAGCTTCAGGAAAGTTCATCCATGAGGCAACTTCC 720  
QY 1354 AAAGAGAGCCCTGACACCTTTCACACCGTTCGAGAAATGCAATTCGCCCATCGCCACCTC 1413  
Db 721 AAAGAGAGAAATGCAATTCGCCCATCGCCACCTTCTAGTCCGCAAGGTGAATCTC----- 775  
QY 1414 CTAGTCGCAAAAGCCAGATCCGAGGACTCGAGCCAGAGCCCATTTCCACTTGACGAGC 1473  
Db 776 ----TTCTTTTCAGCCAGATCCGAGGACTCGAGCCAGAGCCCATTTCCACTTGACGAGC 831  
QY 1474 AGCGCGCTCACCAGGAAATGAGGCTACCACTGAGAGATATGTACATAGGAAATGATAAC 1533  
Db 832 AGCGCGCTCACCAGGAAATGAGGCTACCACTGAGAGATATGTACATAGGAAATGATAAC 891  
QY 1534 GAGAGAACTCTTATCAGGAGACATTTTAAACGCGGATGGCTTTCAGCGTGACCAAT 1593  
Db 892 GAGAGAACTCTTATCAGGAGACATTTTAAACGCGGATGGCTTTCAGCGTGACCAAT 951  
QY 1594 ACAGGCTATATTACGTATACGCCAGATATCTCAACAACTCGCAGACCGAGACGGA 1653  
Db 952 ACAGGCTATATTACGTATACGCCAGATATCTCAACAACTCGCAGACCGAGACGGA 1011  
QY 1654 TTATTCGTCTTTCAAGGAGACACTCATTCCTGCACTGCTTGAACACGGTCCCAACCAAC 1713  
Db 1012 TTATTCGTCTTTCAAGGAGACACTCATTCCTGCACTGCTTGAACACGGTCCCAACCAAC 1071  
QY 1714 ATGCCACATAAGGTGCGACACTGCGCACAGAGTGTCTGATCCACTGGAACGAAACGAG 1773  
Db 1072 ATGCCACATAAGGTGCGACACTGCGCACAGAGTGTCTGATCCACTGGAACGAAACGAG 1131  
QY 1774 AGGATCCCATCTGAAGGACATTCACAAAGATCGCAATGCAATGCGATGCGATGCGGAGGAAACAAAC 1833  
Db 1132 AGGATCCCATCTGAAGGACATTCACAAAGATCGCAATGCGATGCGATGCGGAGGAAACAAAC 1191  
QY 1834 CGAAGCTATTGGGCATCTTCAAGGTGTAA 1863  
Db 1192 CGAAGCTATTGGGCATCTTCAAGGTGTAA 1221

## RESULT 4

US-09-790-399-7  
; Sequence 7, Application US/09790399  
; Patent No. US20020038000A1  
; GENERAL INFORMATION:  
; APPLICANT: Gold, Larry  
; APPLICANT: Tuerk, Craig  
; APPLICANT: Prinslow, David  
; APPLICANT: Smith, Jonathan D.  
; TITLE OF INVENTION: Systematic Polypeptide Evolution by Reverse Translation  
; FILE REFERENCE: NEX02/C1-CON2  
; CURRENT APPLICATION NUMBER: US/09/790,399  
; CURRENT FILING DATE: 2001-02-22  
; PRIOR APPLICATION NUMBER: 09/197,649  
; PRIOR FILING DATE: 1998-11-23  
; PRIOR APPLICATION NUMBER: 07/829,461

```

; PRIOR FILING DATE: 1992-01-31
; PRIOR APPLICATION NUMBER: 07/739,055
; PRIOR FILING DATE: 1991-08-01
; PRIOR APPLICATION NUMBER: 07/561,968
; PRIOR FILING DATE: 1990-08-02
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn ver. 2.0

```

[illegible]

```

RESULT 5
US-10-425-114-7734
; Sequence 7734, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 7734
; LENGTH: 966
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 700727423 FL1

```

US-10-425-114-7734

Query Match	2.78;	Score 57.2;	DB 13;	Length 966;
Best Local Similarity	53.33;	Pred. No. 0.00024;		
Matches 147;	Conservative 0;	Mismatches 123;	Indels 6;	Gaps 1;
Qy	822	TGTATCGCATCTCGACAAAGGAGCTGGAAGACCTCGAACGAGTCGTGATTAATCTCCAGCA	881	
Db	78	TCITTTGGTTCTCGATTTGAAGATTGAAGGCCACAAAGGAATTGAGCCAGCATGTGGAGGA	137	
Qy	882	GCCTTTGGGCATAAATCTATCTGCAGCAGGTTCCGACGAGTTCACAAAGAGGTACGAGATGC	941	
Db	138	AAATGAGGTCACCGTCTTTGGTTGATGAGGAGGAGACGAAGATATACGATCCCGAGGATGA	197	
Qy	942	CCTCATCGACTATCCAAAAAGGTGGATGCGCTCACGGATGAGGAGGACGACGCGATGG	1001	
Db	198	AGTAGTAGACGACGGAGAGGACGATGAAGAC-----GACGACGACGACGATGACGATGA	251	
Qy	1002	CGATGGTCTCGAATCCATTCGCGACGACGAGGACGACGCGTGTAGCTTATAGCTCTGTGGG	1061	
Db	252	CGACGCAATGAAAGACGAGGAACGACGATGAAGATGTCCTCCGACGGCGGTGACGA	311	
Qy	1062	TGATGTTGGCGCAGACTACGAGGACTACCGGATAT	1097	
Db	312	CGACGACGATGTAAGAAGGAAAGAAAGGATGATGT	347	

```

RESULT 6
US-10-424-599-46034
; Sequence 46034, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovacic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 46034
; LENGTH: 1269
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_14156C.1
US-10-424-599-46034

```

	Query Match	2.7%; Score 57.2; DB 13; Length 1269;
	Best Local Similarity 53.3%; Pred. No. 0.00029;	
	Matches 147; Conservative 0; Mismatches 123; Indels 6; Gaps 1;	
Qy	822 TGTATCGCATCTGGACAAAGAGAGCTGAAGACCGCTGAACGGAGTCGTGATATCTCCAGCA 881	
Db	381 TCTTTGGTTCTTGGATTGGAAGATTGAAGGCCACAAAGGAAATTGAGCCAGCATGTGGAGGA 440	
Qy	882 GCGTTTGGGATAAACTATCTGACGACGAGTTCGACGAGTTCCTCAAAGAGATACGAGAATGC 941	
Db	441 AAATGAGGTCAACGGTCTTTGGTTGATGAGGAGGAGGACGAAGAATACGAGTCCCGGAGTGA 500	
Qy	942 CCTCATCGACTATCTCAAAAAAGGTGGATGCGCTCACGGATGAGGAGGACGACACGATGG 1001	
Db	501 AGTAGTAGACGACGGAGAGACGATGNAGAC-----GACGACGACGACGATGACGATGA 554	
Qy	1002 CGATGGTCTCGATTCCATTCCGACGACGAGGACGACGACGTTAGCTTATAGCTCTGTGGGA 1061	
Db	555 CGACGCAATGAAGACGAGGACGACGACGATGAAGATGATGCTCCCGACGCGGTGACGA 614	
Qy	1062 TGATGTTGGCGCAGACTACGAGGACTACCCGATAT 1097	
Db	615 CGACGACGATGAGACCGAAGAAAGAAAGATGATGT 650	



```
; APPLICANT: Roush, Christine L.
; TITLE OF INVENTION: ANTIGENIC PEPTIDES AND ANTIBODIES FOR G PROTEIN-COUPLED RECEPTORS
; FILE REFERENCE: 1920-4-4
; CURRENT APPLICATION NUMBER: US/10/225,567A
; CURRENT FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/257,144
; PRIOR FILING DATE: 2000-12-19
; NUMBER OF SEQ ID NOS: 2292
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 606
; LENGTH: 1915
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-225-567A-606

Query Match          2.5%; Score 53.6; DB 15; Length 1915;
Best Local Similarity 77.4%; Pred. No. 0.0037;
Matches 65; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

Qy 2059 TTGTATATACCTGGTTCGACTCTAAAGCTGAATAAATATATGTAGCTATTAAAAA 2118
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1803 TTTTATAAATCCACCTCTCCAAAAAGCAATAAAAAACAACACTATAAAAAA 1862
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 2119 AAAAAAAAAAAAAAAAAAAAAA 2142
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1863 AAAAAAAAAAAAAAAAAAAAAA 1886
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 11
US-10-424-599-9903
; Sequence 9903, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 9903
; LENGTH: 2190
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(2190)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_108953C.1
US-10-424-599-9903

Query Match          2.5%; Score 53.2; DB 13; Length 2190;
Best Local Similarity 71.4%; Pred. No. 0.0052;
Matches 70; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

Qy 2045 AAATCATATTTATTTTGTATATACCTGTTTCGACTCTAAAAAGTGAATAAATATATGTA 2104
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1928 AAATATATATTTTATATATAAATTTATAAATAAATAAATAAATAAATAAATAA 1987
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 2105 GCTATTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2142
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1988 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2025
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 12
US-10-424-599-82347
; Sequence 82347, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: Birse et al.
```

```
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 82347
; LENGTH: 219
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_45376C.1
US-10-424-599-82347

Query Match          2.5%; Score 52.8; DB 13; Length 219;
Best Local Similarity 71.9%; Pred. No. 0.0013;
Matches 69; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

Qy 2047 ATCATATTTATTTTGTATATACCTGTTTCGACTCTAAAAAGTGAATAAATAATATGTAGC 2106
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 74 ATGTTTTCCTTTCGCTACTCATCTTTTGTCTTTTCAAAAGTGTAAATAAATAATGTTTCA 133
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 2107 TATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2142
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 134 CATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 169
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 13
US-10-424-599-38982/c
; Sequence 38982, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 38982
; LENGTH: 372
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_135200C.1
US-10-424-599-38982

Query Match          2.5%; Score 52.8; DB 13; Length 372;
Best Local Similarity 69.2%; Pred. No. 0.0019;
Matches 72; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

Qy 2039 GGAGGAAATCATATTTATTTTGTATATACCTGTTTCGACTCTAAAAAGTGAATAAATA 2098
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 131 GGGGGAAGGTAATTTTATTTGATGAACCCCTTTCCCTTCAAAAAAAAAA 72
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 2099 TATCTAGCTATTAAAAAATAAATAAATAAATAAATAAATAAATAAATAA 2142
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 71 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 28
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 14
US-10-264-049-287
; Sequence 287, Application US/10264049
; Publication No. US20040005579A1
; GENERAL INFORMATION:
; APPLICANT: Birse et al.
```



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 14, 2004, 18:37:56 ; Search time 8239 Seconds  
(without alignments)  
7785.406 Million cell updates/sec

Title: US-09-813-329-5  
Perfect score: 2148  
Sequence: 1 ggcacggcgacgacgt.....aaaaaaaaaactcgag 2148

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues  
Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:\*  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hic:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_est3:\*  
12: gb\_est4:\*  
13: gb\_est5:\*  
14: em\_estfun:\*  
15: em\_estom:\*  
16: em\_ges\_hum:\*  
17: em\_ges\_inv:\*  
18: em\_ges\_pln:\*  
19: em\_ges\_vrt:\*  
20: em\_ges\_fun:\*  
21: em\_ges\_nam:\*  
22: em\_ges\_mus:\*  
23: em\_ges\_pro:\*  
24: em\_ges\_rod:\*  
25: em\_ges\_phg:\*  
26: em\_ges\_vrl:\*  
27: gb\_ges1:\*  
28: gb\_ges2:\*  
29: gb\_ges2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES			
Result No.	Score	Query Match Length	ID Description
1	686.4	32.0	699 12 BI141711 SD15611.5
2	678	31.6	686 12 BG636194 SD13795.5
3	673.4	31.4	686 12 BI637732 SD19794.5
4	666.8	31.0	688 12 BI638827 SD21218.5

5	632.8	29.5	649	12	BG640917	BG640917 SD12092.5
6	599.2	27.9	604	9	AI260099	AI260099 LP03784.5
7	545.4	25.4	565	12	BI620499	BI620499 RH51659.5
8	500	23.3	508	14	CK135985	CK135985 SD18286.3
9	500	23.3	532	14	CK135525	CK135525 RH51659.3
10	477.4	22.2	495	12	BI636564	BI636564 SD18286.5
11	408.8	19.0	447	12	BI635992	BI635992 SD17545.5
12	129.8	6.0	612	13	BI650278	BI650278 BX560278
13	115.6	5.4	1101	29	CNS017WQ	AL108596 Drosophila
14	65.6	3.1	468	12	BG553005	BG553005 dab82609.
15	62.8	2.9	752	14	CD099026	CD099026 AGENCOURT
16	62.8	2.9	923	13	BU771970	BU771970 SJBELA02
17	62	2.9	697	13	BQ524571	BQ524571 NISC no05
18	61.8	2.9	935	14	CK017088	CK017088 AGENCOURT
19	61	2.8	887	13	BQ731479	BQ731479 AGENCOURT
20	59.6	2.8	1942	13	BU771792	BU771792 SJEJA08
21	59.4	2.8	1091	13	BU766419	BU766419 SJAERA02
22	59.2	2.8	863	14	CD752688	CD752688 AGENCOURT
23	59	2.7	866	12	BM015504	BM015504 603641716
24	59	2.7	887	12	BG166263	BG166263 602345444
25	58.6	2.7	2162	13	BU766490	BU766490 SJAER03
26	58.2	2.7	811	14	CF147545	CF147545 AGENCOURT
27	57.6	2.7	697	10	AW134038	AW134038 fl14903.y
28	57	2.7	198	14	CF209778	CF209778 CAB20005
29	57	2.7	299	14	CB366950	CB366950 OI305 oOk
30	57	2.7	664	14	CB066301	CB066301 PVBE12D11
31	57	2.7	795	28	AZ528485	AZ528485 ENTCM64TP
32	57	2.7	823	28	AZ676218	AZ676218 ENTKE36TR
33	57	2.7	828	13	BW249956	BW249956 BW249956
34	57	2.7	843	28	BH139532	BH139532 ENTNG88TF
35	57	2.7	1124	13	BQ421045	BQ421045 AGENCOURT
36	56.6	2.6	933	13	BX706375	BX706375 BX706375
37	56.4	2.6	180	12	BM307698	BM307698 eak32g05.
38	56.4	2.6	501	13	C84767	C84767 C84767 Dict
39	56.2	2.6	681	29	CNS02EOD	AL193990 Tetradon
40	56	2.6	645	12	BM165350	BM165350 EST567873
41	56	2.6	647	12	BM163120	BM163120 EST565643
42	56	2.6	671	12	BM160252	BM160252 EST562775
43	56	2.6	694	12	BM168242	BM168242 EST570765
44	56	2.6	712	12	BM394933	BM394933 50072-2-6
45	56	2.6	717	12	BM160500	BM160500 EST563023

ALIGNMENTS

RESULT 1  
BI141711  
LOCUS  
DEFINITION SD15611.5prime SD Drosophila melanogaster Schneider L2 cell culture  
EST 03-JUL-2001  
BI141711  
ACCSSION BI141711  
VERSION BI141711  
KEYWORDS BI141711.1 GI:14594155  
SOURCE EST.  
ORGANISM Drosophila melanogaster (fruit fly)  
Drosophila melanogaster  
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
Ephydroidea; Drosophilidae; Drosophila.  
REFERENCE 1 (bases 1 to 699)  
AUTHORS Harvey,D., Brokstein,P., Hong,L., Evans-Holm,M., Su,C., Tsang,G.,  
Lewis,S. and Rubin,G.M.  
TITLE BQGP/HMI Drosophila EST Project  
JOURNAL Unpublished (2001)  
COMMENT Contact: Stapleton, M.  
BDGP  
Lawrence Berkeley National Lab  
One Cyclotron Rd. Berkeley, CA 94720, USA  
Fax: 510 486 6798  
Email: <http://www.fruitfly.org/EST>, [est@fruitfly.berkeley.edu](mailto:est@fruitfly.berkeley.edu)  
hit genomic AE003831: arm:2R [4876906,5137815]  
estimated-cyto:46B13-46E4: 04/13/2001  
Plate: SD.156 row: A column: 11

High quality sequence stop: 698.

FEATURES  
source

1. .699  
Location/Qualifiers  
/organism="Drosophila melanogaster"  
/mol\_type="mRNA"  
/db\_xref="taxon:7227"  
/clone="SD15611"  
/lab\_host="DH5-alpha"  
/clone\_lib="SD Drosophila melanogaster Schneider L2 cell culture pot2"  
/note="Vector: pot2; Site 1: EcoRI; Site 2: XhoI; Sized fractionated cDNAs were directly ligated into pot2. Plasmid cDNA library."

ORIGIN

Query Match 32.0%; Score 686.4; DB 12; Length 699;  
Best Local Similarity 99.7%; Pred. No. 1.6e-76;  
Matches 698; Conservative 0; Mismatches 1; Indels 1; Gaps 1;  
QY 10 CGAAGCGAGCTTTAAAGTGAGAAAAGAACCGGTAAATCAGAGATCCCAAGACGCGGT 69  
Db 1 CGAAGCGAGCTTTAAAGTGAGAAAAGAACCGGTAAATCAGAGATCCCAAGACGCGGT 60  
QY 70 GCGTGCATGATAGCGAAGAAAAGCTATCGTTTTCAGTTAACTTACCAAGATTGA 129  
Db 61 GCGTGCATGATAGCGAAGAAAAGCTATCGTTTTCAGTTAACTTACCAAGATTGA 120  
QY 130 ATTTCCGCATCGGCAAAATTAATAAATACATAGTCAACTCGTCCACTGTGTGTGT 189  
Db 121 ATTTCCGCATCGGCAAAATTAATAAATACATAGTCAACTCGTCCACTGTGTGTGT 180  
QY 190 GTTTTTTTTTTTTTTTTTTTTTTTCGCTGTGCTTTTATCGCAAAACAAGAACTGATAAAC 249  
Db 181 G-TTTTTTTTTTTTTTTTTTTTTTTCGCTGTGCTTTTATCGCAAAACAAGAACTGATAAAC 239  
QY 250 TAGAAATATCTTGAGAACTGTGTTTCGCGCTTTTCTTTTGTGTAATGCGGGA 309  
Db 240 TAGAAATATCTTGAGAACTGTGTTTCGCGCTTTTCTTTTGTGTAATGCGGGA 299  
QY 310 AGAGAAAACAAGCAGTAGACAAAACAAGTGTGTTGTAATACATCTGAAAGGGCACCATC 369  
Db 300 AGAGAAAACAAGCAGTAGACAAAACAAGTGTGTTGTAATACATCTGAAAGGGCACCATC 359  
QY 370 AGCAGCCCGAGGGTTTATCTATAGATGTCGAGCTTATCATCTCATGCTGTCTGTGA 429  
Db 360 AGCAGCCCGAGGGTTTATCTATAGATGTCGAGCTTATCATCTCATGCTGTCTGTGA 419  
QY 430 GTTTGTTCTGTGCTCGTGTAGTATCTTAAATACATAGAGTGTGTTTCAATAAGTGCG 489  
Db 420 GTTTGTTCTGTGCTCGTGTAGTATCTTAAATACATAGAGTGTGTTTCAATAAGTGCG 479  
QY 490 ACAAGCTCGATTGGAAACAGCTGTCAGTGCCTTTCAGTGGGTGGGCAAGATCGTCATC 549  
Db 480 ACAAGCTCGATTGGAAACAGCTGTCAGTGCCTTTCAGTGGGTGGGCAAGATCGTCATC 539  
QY 550 ATCATATCTGTCGTCATTATCAACAGAAATCAGCATCAGCATCGAGTGTGTTTATAGCCCA 609  
Db 540 ATCATATCTGTCGTCATTATCAACAGAAATCAGCATCAGCATCGAGTGTGTTTATAGCCCA 599  
QY 610 TAAGATCCCAAGTGTTCATCAATTTATGACTGCGGAGACCTCAAGCGTTTATAGCCCA 669  
Db 600 TAAGATCCCAAGTGTTCATCAATTTATGACTGCGGAGACCTCAAGCGTTTATAGCCCA 659  
QY 670 ACAGTCCCAACGATGATGTTTTCGCGCAAAAGCGACCA 709  
Db 660 ACAGTCCCAACGATGATGTTTTCGCGCAAAAGCGACCA 699

RESULT 2  
BG636194 686 bp mRNA linear EST 23-APR-2001  
LOCUS SD13795.5prime SD Drosophila melanogaster Schneider L2 cell culture  
DEFINITION pot2 Drosophila melanogaster cDNA clone SD13795 5 similar to

CG12919: FBan0012919 located on: 2R 46E1-46E1;: 04/13/2001, mRNA sequence.  
BG636194  
BG636194.1 GI:13763731  
EST.  
Drosophila melanogaster (fruit fly)  
Drosophila melanogaster  
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.  
1 (bases 1 to 686)  
Harvey,D., Brokstein,P., Hong,L., Evans-Holm,M., Su,C., Tsang,G., Lewis,S. and Rubin,G.M.  
BDGP/HHMI Drosophila EST Project  
Unpublished (2001)  
Contact: Stapleton, M.  
BDGP  
Lawrence Berkeley National Lab  
One Cyclotron Rd, Berkeley, CA 94720, USA  
Fax: 510 486 6798  
Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu  
hit genomic AB003831: arm:2R [4876906,5137815]  
estimated-cyto:46B13-46B4: 04/13/2001  
Plate: SD.137 row: H column: 11  
High quality sequence stop: 685.  
Location/Qualifiers  
1. .686  
/organism="Drosophila melanogaster"  
/mol\_type="mRNA"  
/db\_xref="taxon:7227"  
/clone="SD13795"  
/lab\_host="DH5-alpha"  
/clone\_lib="SD Drosophila melanogaster Schneider L2 cell culture pot2"  
/note="Vector: pot2; Site 1: EcoRI; Site 2: XhoI; Sized fractionated cDNAs were directly ligated into pot2. Plasmid cDNA library."

ORIGIN

Query Match 31.6%; Score 678; DB 12; Length 686;  
Best Local Similarity 99.3%; Pred. No. 1.8e-75;  
Matches 681; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
QY 1401 CCATCGCCACCTCTCTAGTCCGCAAGCCAGATCCGAGACTCGAGGCCAGAGCCCATTT 1460  
Db 1 CCATCGCCACCTCTCTAGTCCGCAAGCCAGATCCGAGACTCGAGGCCAGAGCCCATTT 60  
QY 1461 CCATCTGAGCAGCAGCGCGCTCACCAGGAAGTATGGCTTACCATGGAGATATGTACAT 1520  
Db 61 CCACCTGAGCAGCAGCGCGCTCACCAGGAAGTATGGCTTACCATGGAGATATGTACAT 120  
QY 1521 AGGAAATGATACGAGAGAACTCTTATCAGGACACTTTCAAACGCGCGATGCGTCTT 1580  
Db 121 AGGAAATGATACGAGAGAACTCTTATCAGGACACTTTCAAACGCGCGATGCGTCTT 180  
QY 1581 GACGTGACCAATACAGGCTTATATAGTATAGCCAGATATGCTTCCATTCCTCAGTGTGTAACAC 1640  
Db 181 GACGTGACCAATACAGGCTTATATAGTATAGCCAGATATGCTTCCATTCCTCAGTGTGTAACAC 240  
QY 1641 CGACAGAACGGAATTTATCGTCTTTTCAAGGAGACACTCCATTCCTCAGTGTGTAACAC 1700  
Db 241 CGACAGAACGGAATTTATCGTCTTTTCAAGGAGACACTCCATTCCTCAGTGTGTAACAC 300  
QY 1701 GGTGCCCAACCATGCCATAGTGTGCACACTGCCACAGAGTGGTCTGATCCACT 1760  
Db 301 GGTGCCCAACCATGCCATAGTGTGCACACTGCCACAGAGTGGTCTGATCCACT 360  
QY 1761 GGAAACGAAACGAGGATCCATCTGAAGGACATTCACAAACGATCGCAATCGAGTCTTGGC 1820  
Db 361 GGAAACGAAACGAGGATCCATCTGAAGGACATTCACAAACGATCGCAATCGAGTCTTGGC 420  
QY 1821 GGAGGAAACACCGAAGCTTACTTTGGCATCTTCAAGGTGTAATGGAGAGATTATCCC 1880  
|||||



```

Db      421 GGAGGAAACACCGAAGCTACTTTGGCATCTTCAAGGTGTAATAATGGAGAGATTATCC 480
Qy      1881 CGGTGAGAGATGAATACCAAGTTTAAAGCTTTTGTCCCGGAGTGTCTCGTGAATCGGAT 1940
Db      481 CGGTGAGAGATGAATACCAAGTTTAAAGCTTTTGTCCCGGAGTGTCTCGTGAATCGGAT 540
Qy      1941 TCATCGCCAGCGTGAATCCATAGTTTGGTAGTACCTAGTCTTAGTCACTCCAAACCTAAT 2000
Db      541 TCATCGCCAGCGTGAATCCATAGTTTGGTAGTACCTAGTCTTAGTCACTCCAAACCTAAT 600
Qy      2001 CTCATCGGATCGTCATCTGATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCAT 2060
Db      601 CTCATCGGATCGTCATCTGATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCAT 660
Qy      2061 GTATATACCTGTCGACTCTAAAAAG 2086
Db      661 GTATATACCTGTCGACTCTAAAAAG 686

RESULT 3
LOCUS   BI637732                686 bp      mRNA      linear      EST 10-SEP-2001
DEFINITION SD19794.Sprime SD Drosophila melanogaster cDNA clone SD19794 5, mRNA sequence.
ACCESSION BI637732
VERSION   1
KEYWORDS  Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
SOURCE    Ephydroidea; Drosophilidae; Drosophila.
ORGANISM  Drosophila melanogaster (fruit fly)
AUTHORS   Harvey,D., Brokstein,P., Hong,L., Evans-Holm,M., Su,C., Tsang,G.,
          Lewis,S. and Rubin,G.M.
TITLE     BDGP/HMI Drosophila EST Project
JOURNAL   Unpublished (2001)
COMMENT   Contact: Stapleton, M.
          BDGP
          Lawrence Berkeley National Lab
          One Cyclotron Rd, Berkeley, CA 94720, USA
          Fax: 510 486 6798
          Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
          hit genomic AE003831: arm:2R [4876906,5137815]
          estimated-cyto:46B13-46E4: 05/19/2001
          Plate: SD.197 row: H column: 10
          High quality sequence stop: 584.
FEATURES
         source
         1..686
         /organism="Drosophila melanogaster"
         /mol_type="mRNA"
         /db_xref="taxon:7227"
         /clone="SD19794"
         /lab_hosts="DHS-alpha"
         /clone_lib="SD Drosophila melanogaster Schneider L2 cell
         culture pOT2"
         /note="vector: pOT2; Site.1: EcoRI; Site 2: XhoI; Sized
         fractionated cDNAs were directly ligated into pOT2.
         Plasmid cDNA library."

ORIGIN
Query Match 31.4%; Score 673.4; DB 12; Length 686;
Beet Local Similarity 99.7%; Pred. No. 6.6e-75;
Matches 685; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

Qy      10 CGAACGACGTTTAAAGTGAGAAAAGAACCGGTAAATCAGAGATCCCAAGCAAGCGCGT 69
Db      1 CGAACGACGTTTAAAGTGAGAAAAGAACCGGTAAATCAGAGATCCCAAGCAAGCGCGT 60
Qy      70 GCGTGCATGATAGCGAAGAAAAGCTATCCGTTTCAGTTAACTACTTACCAAGATTGA 129
Db      61 GCGTGCATGATAGCGAAGAAAAGCTATCCGTTTCAGTTAACTACTTACCAAGATCGA 120

```

```

Qy      130 ATTTGCGCATCGGCGAAATTAATAAAATACATAAGTGCACCTCGTCCACTGTGTGTGT 189
Db      121 ATTTGCGCATCGGCGAAATTAATAAAATACATAAGTGCACCTCGTCCACTGTGTGTGT 180
Qy      190 GTTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 249
Db      181 G-TTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 239
Qy      250 TAGAAAAATATCTTGGAGAACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 309
Db      240 TAGAAAAATATCTTGGAGAACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 299
Qy      310 AGAGAAAAACAACAGCAGTAGACAAAAACAAGTGTGTGTGTGTGTGTGTGTGTGTGTGT 369
Db      300 AGAGAAAAACAACAGCAGTAGACAAAAACAAGTGTGTGTGTGTGTGTGTGTGTGTGTGT 359
Qy      370 AGAGCCCCGAGGGGTTTATCTATATAGATGTCCAGCTTATCATCTCATCTGTCTGTGA 429
Db      360 AGAGCCCCGAGGGGTTTATCTATATAGATGTCCAGCTTATCATCTCATCTGTCTGTGA 419
Qy      430 GGTGTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 489
Db      420 GGTGTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 479
Qy      490 ACAAAGCTCGATTGGAAACAGCTGTGAGTGCCCTTGAGTGGGTGGGCAAGATCGTCAATC 549
Db      480 ACAAAGCTCGATTGGAAACAGCTGTGAGTGCCCTTGAGTGGGTGGGCAAGATCGTCAATC 539
Qy      550 ATCATCATCTGTCTATTATCAACAGATCAGCATCAGCATCTGGAGGCCCGGTCTCTC 609
Db      540 ATCATCATCTGTCTATTATCAACAGATCAGCATCAGCATCTGGAGGCCCGGTCTCTC 599
Qy      610 TAAGATCCCCAGTGTTCATCAATTATGACTGCCGAGACCTCAAGCCGTTTATAACGCCA 669
Db      600 TAAGATCCCCAGTGTTCATCAATTATGACTGCCGAGACCTCAAGCCGTTTATAACGCCA 659
Qy      670 ACGAGTCCCAACGATGATGTTTCCG 696
Db      660 ACGAGTCCCAACGATGATGTTTCCG 686

RESULT 4
LOCUS   BI638827                688 bp      mRNA      linear      EST 10-SEP-2001
DEFINITION SD21218.Sprime SD Drosophila melanogaster Schneider L2 cell culture
          pOT2 Drosophila melanogaster cDNA clone SD21218 5, mRNA sequence.
ACCESSION BI638827
VERSION   1
KEYWORDS  Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
SOURCE    Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
          Ephydroidea; Drosophilidae; Drosophila.
ORGANISM  Drosophila melanogaster (fruit fly)
AUTHORS   Harvey,D., Brokstein,P., Hong,L., Evans-Holm,M., Su,C., Tsang,G.,
          Lewis,S. and Rubin,G.M.
TITLE     BDGP/HMI Drosophila EST Project
JOURNAL   Unpublished (2001)
COMMENT   Contact: Stapleton, M.
          BDGP
          Lawrence Berkeley National Lab
          One Cyclotron Rd, Berkeley, CA 94720, USA
          Fax: 510 486 6798
          Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
          hit genomic AE003831: arm:2R [4876906,5137815]
          estimated-cyto:46B13-46E4: 05/19/2001
          Plate: SD.212 row: B column: 6
          High quality sequence stop: 602.
FEATURES
         source
         1..688
         /organism="Drosophila melanogaster"
         /mol_type="mRNA"

```

```
/db_xref="taxon:7227"  
/clone="SD21218"  
/lab_host="DH5-alpha"  
/clone_lib="SD Drosophila melanogaster Schneider L2 cell  
culture pot2"  
/note="Vector: pot2; Site 1: EcoRI; Site 2: XhoI; Sized  
fractionated cDNAs were directly ligated into pot2.  
Plasmid cDNA library."
```

## ORIGIN

Query Match	31.0%;	Score 666.8;	DB 12;
Best Local Similarity	99.4%;	Pred. No. 4.4e-74;	
Matches 680;	Conservative	0;	Mismatches 2;
			Indels 2;
			Gaps 1;

QY	5	CGAGCGGAACGGGACGTTTAAAGTGAAGAAAGAAACCGGTAAATCAGAGATCCCAAGCAAG	64
Db	7	CGAAGCGAACGGGACGTTTAAAGTGAAGAAAGAAACCGGTAAATCAGAGATCCCAAGCAAG	66
QY	65	CGCGTGCCTGCATGATAGCGAAGAAAGAAAGCTATCCGTTTCAGTTAACTTACTTAACCAAG	124
Db	67	CGCGTGCCTGCATGATAGCGAAGAAAGAAAGCTATCCGTTTCAGTTAACTTACTTAACCAAG	126
QY	125	ATTGAATTTCCGCATCGGCGCAAAATTACTTAAATAATACATAAGTGCAAACTCGTCCACTGTGT	184
Db	127	ATCGAAATTTCCGCATCGGCGCAAAATTACTTAAATAATACATAAGTGCAAACTCGTCCACTGTGT	186
QY	185	GTTGTGTTTTTTTTTTTTTTTTTTTTTTTTGGTTTTTCGCTGCGCTTATCGCAAAACAGAACTGAT	244
Db	187	GTTGTG--TTTTTTTTTTTTTTTTTTTTTTTTGGTTTTTCGCTGCGCTTATCGCAAAACAGAACTGAT	244
QY	245	AAAACTAGAAAATCTTTGGAACACTTGTTTTCGCGCTTTTCTTTTCGCTAAATGCGCGATC	304
Db	245	AAAACTAGAAAATCTTTGGAACACTTGTTTTCGCGCTTTTCTTTTCGCTAAATGCGCGATC	304
QY	305	GCGGAAGAGAAAACACGACGATAGACAAAACAAGTGTGGTAATACAATCTGAAAAAGGCA	364
Db	305	GCGGAAGAGAAAACACGACGATAGACAAAACAAGTGTGGTAATACAATCTGAAAAAGGCA	364
QY	365	CCATCAGCAGCCGAGGGGTTTACTATATAGATGTCGACGCTTATCATCTCAATGCTGTC	424
Db	365	CCATCAGCAGCCGAGGGGTTTATCATATATAGATGTCGACGCTTATCATCTCAATGCTGTC	424
QY	425	TGTGAGGTTGTTCTGTGTGTCGTGTGTAGTATCTTAAATACATAGAGTGTGTTCATATAAA	484
Db	425	TGTGAGGTTGTTCTGTGTGTCGTGTGTAGTATCTTAAATACATAGAGTGTGTTCATATAAA	484
QY	485	GTGCGACAAAGCTCGATTGGAACACAGCTGTCGAGTGCCCTTGAGTCGGTGGGCAAGATCG	544
Db	485	GTGCGACAAAGCTCGATTGGAACACAGCTGTCGAGTGCCCTTGAGTCGGTGGGCAAGATCG	544
QY	545	TCATCATCATCTGTGCTCATTTCAACAGAAATCAGCATCAGCATCTCGAGAGCCCGCGT	604
Db	545	TCATCATCATCTGTGCTCATTTATCAACAGAAATCAGCATCAGCATCTCGAGAGCCCGCGT	604
QY	605	TGCTCTAAGATCCCGAGTGTTCAATCAATATGACTGCCGAGACCCCTCAAGCCGTTTATAA	664
Db	605	TGCTCTAAGATCCCGAGTGTTCAATCAATATGACTGCCGAGACCCCTCAAGCCGTTTATAA	664
QY	665	CGCCAAACGAGTGCCAAACGATGATG	688
Db	665	CGCCAAACGAGTGCCAAACGATGATG	688

RESULT 5	649 bp	linear	EST 23-APR-2001
BG640917	SD	Drosophila	melanogaster
LOCUS	SD12092.5	prime	Schneider L2 cell culture
DEFINITION	POT2	Drosophila	melanogaster cDNA clone SD12092 5, mRNA sequence.
ACCESSION	BG640917		
VERSION	BG640917.1	GI:13772843	
KEYWORDS	EST.		
SOURCE	Drosophila	melanogaster	(fruit fly)
ORGANISM	Drosophila	melanogaster	

Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
Ephydroidea; Drosophilidae; Drosophila.  
1 (bases 1 to 649)  
Harvey, D., Brokstein, P., Hong, L., Evans-Holm, M., Su, C., Tsang, G.,  
Lewis, S. and Rubin, G.M.  
BDGP/HHMI Drosophila EST Project  
Unpublished (2001)  
Contact: Stapleton, M.  
BDGP

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1	...	...	...	...
2	...	...	...	...
3	...	...	...	...
4	...	...	...	...
5	...	...	...	...
6	...	...	...	...
7	...	...	...	...
8	...	...	...	...
9	...	...	...	...
10	...	...	...	...
11	...	...	...	...
12	...	...	...	...
13	...	...	...	...
14	...	...	...	...
15	...	...	...	...
16	...	...	...	...
17	...	...	...	...
18	...	...	...	...
19	...	...	...	...
20	...	...	...	...
21	...	...	...	...
22	...	...	...	...
23	...	...	...	...
24	...	...	...	...
25	...	...	...	...
26	...	...	...	...
27	...	...	...	...
28	...	...	...	...
29	...	...	...	...
30	...	...	...	...
31	...	...	...	...
32	...	...	...	...
33	...	...	...	...
34	...	...	...	...
35	...	...	...	...
36	...	...	...	...
37	...	...	...	...
38	...	...	...	...
39	...	...	...	...
40	...	...	...	...
41	...	...	...	...
42	...	...	...	...
43	...	...	...	...
44	...	...	...	...
45	...	...	...	...
46	...	...	...	...
47	...	...	...	...
48	...	...	...	...
49	...	...	...	...
50	...	...	...	...
51	...	...	...	...
52	...	...	...	...
53	...	...	...	...
54	...	...	...	...
55	...	...	...	...
56	...	...	...	...
57	...	...	...	...
58	...	...	...	...
59	...	...	...	...
60	...	...	...	...
61	...	...	...	...
62	...	...	...	...
63	...	...	...	...
64	...	...	...	...
65	...	...	...	...
66	...	...	...	...
67	...	...	...	...
68	...	...	...	...
69	...	...	...	...
70	...	...	...	...
71	...	...	...	...
72	...	...	...	...
73	...	...	...	...
74	...	...	...	...
75	...	...	...	...
76	...	...	...	...
77	...	...	...	...
78	...	...	...	...
79	...	...	...	...
80	...	...	...	...
81	...	...	...	...
82	...	...	...	...
83	...	...	...	...
84	...	...	...	...
85	...	...	...	...
86	...	...	...	...
87	...	...	...	...
88	...	...	...	...
89	...	...	...	...
90	...	...	...	...
91	...	...	...	...
92	...	...	...	...
93	...	...	...	...
94	...	...	...	...
95	...	...	...	...
96	...	...	...	...
97	...	...	...	...
98	...	...	...	...
99	...	...	...	...
100	...	...	...	...

Lawrence Berkeley National Lab  
One Cyclotron Rd, Berkeley, CA 94720, USA  
Fax: 510 486 6798  
Email: [http://www.fruitfly.org/EST\\_estofruitfly.berkeley.edu](http://www.fruitfly.org/EST_estofruitfly.berkeley.edu)  
hit genomic A5003831: arm:2R [4876906,5137815]  
estimated-cyto:4E13-4E4: 04/13/2001  
Plate: SD.120 row: H column: 8  
High quality sequence stop: 597.

FEATURES  
source

```

1. 649
2. 650
3. 651
4. 652
5. 653
6. 654
7. 655
8. 656
9. 657
10. 658
11. 659
12. 660
13. 661
14. 662
15. 663
16. 664
17. 665
18. 666
19. 667
20. 668
21. 669
22. 670
23. 671
24. 672
25. 673
26. 674
27. 675
28. 676
29. 677
30. 678
31. 679
32. 680
33. 681
34. 682
35. 683
36. 684
37. 685
38. 686
39. 687
40. 688
41. 689
42. 690
43. 691
44. 692
45. 693
46. 694
47. 695
48. 696
49. 697
50. 698
51. 699
52. 700
53. 701
54. 702
55. 703
56. 704
57. 705
58. 706
59. 707
60. 708
61. 709
62. 710
63. 711
64. 712
65. 713
66. 714
67. 715
68. 716
69. 717
70. 718
71. 719
72. 720
73. 721
74. 722
75. 723
76. 724
77. 725
78. 726
79. 727
80. 728
81. 729
82. 730
83. 731
84. 732
85. 733
86. 734
87. 735
88. 736
89. 737
90. 738
91. 739
92. 740
93. 741
94. 742
95. 743
96. 744
97. 745
98. 746
99. 747
100. 748
101. 749
102. 750
103. 751
104. 752
105. 753
106. 754
107. 755
108. 756
109. 757
110. 758
111. 759
112. 760
113. 761
114. 762
115. 763
116. 764
117. 765
118. 766
119. 767
120. 768
121. 769
122. 770
123. 771
124. 772
125. 773
126. 774
127. 775
128. 776
129. 777
130. 778
131. 779
132. 780
133. 781
134. 782
135. 783
136. 784
137. 785
138. 786
139. 787
140. 788
141. 789
142. 790
143. 791
144. 792
145. 793
146. 794
147. 795
148. 796
149. 797
150. 798
151. 799
152. 800
153. 801
154. 802
155. 803
156. 804
157. 805
158. 806
159. 807
160. 808
161. 809
162. 810
163. 811
164. 812
165. 813
166. 814
167. 815
168. 816
169. 817
170. 818
171. 819
172. 820
173. 821
174. 822
175. 823
176. 824
177. 825
178. 826
179. 827
180. 828
181. 829
182. 830
183. 831
184. 832
185. 833
186. 834
187. 835
188. 836
189. 837
190. 838
191. 839
192. 840
193. 841
194. 842
195. 843
196. 844
197. 845
198. 846
199. 847
200. 848
201. 849
202. 850
203. 851
204. 852
205. 853
206. 854
207. 855
208. 856
209. 857
210. 858
211. 859
212. 860
213. 861
214. 862
215. 863
216. 864
217. 865
218. 866
219. 867
220. 868
221. 869
222. 870
223. 871
224. 872
225. 873
226. 874
227. 875
228. 876
229. 877
230. 878
231. 879
232. 880
233. 881
234. 882
235. 883
236. 884
237. 885
238. 886
239. 887
240. 888
241. 889
242. 890
243. 891
244. 892
245. 893
246. 894
247. 895
248. 896
249. 897
250. 898
251. 899
252. 900
253. 901
254. 902
255. 903
256. 904
257. 905
258. 906
259. 907
260. 908
261. 909
262. 910
263. 911
264. 912
265. 913
266. 914
267. 915
268. 916
269. 917
270. 918
271. 919
272. 920
273. 921
274. 922
275. 923
276. 924
277. 925
278. 926
279. 927
280. 928
281. 929
282. 930
283. 931
284. 932
285. 933
286. 934
287. 935
288. 936
289. 937
290. 938
291. 939
292. 940
293. 941
294. 942
295. 943
296. 944
297. 945
298. 946
299. 947
300. 948
301. 949
302. 950
303. 951
304. 952
305. 953
306. 954
307. 955
308. 956
309. 957
310. 958
311. 959
312. 960
313. 961
314. 962
315. 963
316. 964
317. 965
318. 966
319. 967
320. 968
321. 969
322. 970
323. 971
324. 972
325. 973
326. 974
327. 975
328. 976
329. 977
330. 978
331. 979
332. 980
333. 981
334. 982
335. 983
336. 984
337. 985
338. 986
339. 987
340. 988
341. 989
342. 990
343. 991
344. 992
345. 993
346. 994
347. 995
348. 996
349. 997
350. 998
351. 999
352. 1000
353. 1001
354. 1002
355. 1003
356. 1004
357. 1005
358. 1006
359. 1007
360. 1008
361. 1009
362. 1010
363. 1011
364. 1012
365. 1013
366. 1014
367. 1015
368. 1016
369. 1017
370. 1018
371. 1019
372. 1020
373. 1021
374. 1022
375. 1023
376. 1024
377. 1025
378. 1026
379. 1027
380. 1028
381. 1029
382. 1030
383. 1031
384. 1032
385. 1033
386. 1034
387. 1035
388. 1036
389. 1037
390. 1038
391. 1039
392. 1040
393. 1041
394. 1042
395. 1043
396. 1044
397. 1045
398. 1046
399. 1047
400. 1048
401. 1049
402. 1050
403. 1051
404. 1052
405. 1053
406. 1054
407. 1055
408. 1056
409. 1057
410. 1058
411. 1059
412. 1060
413. 1061
414. 1062
415. 1063
416. 1064
417. 1065
418. 1066
419. 1067
420. 1068
421. 1069
422. 1070
423. 1071
424. 1072
425. 1073
426. 1074
427. 1075
428. 1076
429. 1077
430. 1078
431. 1079
432. 1080
433. 1081
434. 1082
435. 1083
436. 1084
437. 1085
438. 1086
439. 1087
440. 1088
441. 1089
442. 1090
443. 1091
444. 1092
445. 1093
446. 1094
447. 1095
448. 1096
449. 1097
450. 1098
451. 1099
452. 1100
453. 1101
454. 1102
455. 1103
45
```

## ORIGIN

Query Match 29.5%; Score 632.8; DB 12; Length 649;  
Best Local Similarity 99.4%; Pred. No. 7.9e-70;  
Matches 646; Conservative 0; Mismatches 2; Indels 2; Gaps 1;

QY	5	CGAGCGAACGGAGCTTTAAAGTCAGAAAAAAGAAACCGGTAAATCAGAGATCCCAAGCAAG	64
DB	2	CGAAGCGAACGGAGCTTTAAAGTCAGAAAAAAGAAACCGGTAAATCAGAGATCCCAAGCAAG	61
QY	65	CGCGTCGGTGCATGATACGAGAGAAAAAGCTATCCGTTTCAGTTAACTACTTTACCAAG	124
DB	62	CGCGTCGGTGCATGATACGAGAGAAAAAAGCTATCCGTTTCAGTTAACTACTTTACCAAG	121
QY	125	ATTGAATTTTCGCCATTCGGGCAAAATTACTAAAAATACATAAGTCGCAACTCGTCCACTGTGT	184
DB	122	ATCGAAATTTTCGCCATTCGGGCAAAATTACTAAAAATACATAAGTCGCAACTCGTCCACTGTGT	181
QY	185	GTTCTGTTTTTTTTTTTTTTTTTTTTTTTTCGCTGTGCTTTATTCGCAAAACAAGACTGAT	244
DB	182	GTGTGTG--TTTTTTTTTTTTTTTTTTTTTTTCGCTGTGCTTTATTCGCAAAACAAGACTGAT	239
QY	245	AAAACTAGAAAAATATCTGTGAAACCTGTGTTTCGCGCTTTTCTTTTGCTAAATGCCGATC	304
DB	240	AAAACTAGAAAAATATCTGTGAAACCTGTGTTTCGCGCTTTTCTTTTGCTAAATGCCGATC	299
QY	305	CGCGAAGAGAAAAACAAGCAGTAGAGCAAAACAAGCTGTGGTAATACATCTGAAACGGCA	364
DB	300	CGCGAAGAGAAAAACAAGCAGTAGAGCAAAACAAGCTGTGGTAATACATCTGAAACGGCA	359
QY	365	CCATCTAGAGCCCGAGGGGTTTATCTATPAGATGTCGACGTTTATCATCTCATGTGTC	424
DB	360	CCATCTAGAGCCCGAGGGGTTTATCTATPAGATGTCGACGTTTATCATCTCATGTGTC	419
QY	425	TGTGAGGTTTGTCTGTGTGCTCGTGTAGTATCTTAAATACATAGAGTGTTTCATATATA	484
DB	420	TGTGAGGTTTGTCTGTGTGCTCGTGTAGTATCTTAAATACATAGAGTGTTTCATATATA	479
QY	485	GTCCGACAAAGTCGATTTGGAAACAGCTGTCGAGTGCCCTTGAGTGGGTGGGCAAGATCG	544
DB	480	GTCCGACAAAGTCGATTTGGAAACAGCTGTCGAGTGCCCTTGAGTGGGTGGGCAAGATCG	539
QY	545	TCATCATCATCATCGTCGTCGTCATTATCAACAGAAATCAGCATCAGCATCTCGAGGCCCGGT	604

```
Db      540 TCATCATCATCGTCGTCATTATCAACAGAATCAGCATCAGCATCTCGGAGCCCGGT 599
Qy      605 TGCTCTAGATCCCGAGGTTCATCAATATGACTGCGGAGACCCCTCAAG 654
Db      600 TGCTCTAAGATCCCGAGGTTCATCAATATGACTGCGGAGACCCCTCAAG 649

RESULT 6
LOCUS   AI260099                604 bp mRNA linear EST 19-APR-2001
DEFINITION LP03784.5prime LP Drosophila melanogaster larval-early pupal pot2
Drosophila melanogaster cDNA clone LP03784 5prime, mRNA sequence.
ACCESSION AI260099
VERSION   AI260099.1 GI:3867624
KEYWORDS EST.
SOURCE   Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 604)
Harvey,D., Brokstein,P., Hong,L., Evans-Holm,M., Su,C., Tsang,G.,
Lewis,S. and Rubin,G.M.
BDGP/HHMI Drosophila EST Project
Unpublished (2001)
Contact: Stapleton, M.
BDGP

Lawrence Berkeley National Lab
One Cyclotron Rd, Berkeley, CA 94720, USA
Fax: 510 486 6798
Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
hit genomic sequence D905033
Plate: 37 row: G column: 12
High quality sequence stop: 421.
Location/Qualifiers
1..604
/organism="Drosophila melanogaster"
/mol_type="mRNA"
/db_xref="taxon:7227"
/clone="LP03784"
/sex="male and female"
/dev_stage="larvae-pupae"
/lab_host="DHS-alpha"
/clone_lib="LP Drosophila melanogaster larval-early pupal
pot2"
/notes="Organ: whole body; Vector: pot2; Site 1: EcoRI;
Site 2: XhoI; Sized fractionated cDNAs were directly
ligated into pot2. Plasmid cDNA library.
"

FEATURES
source
1..604

ORIGIN
Query Match 27.9%; Score 599.2; DB 9; Length 604;
Best Local Similarity 99.5%; Pred. NO. 1.3e-65;
Matches 601; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 10 CGAACGGACGTTTAAAGTGAGAAAAGAAACCGGTAAATCAGAGATCCCAAGCAAGCGCGT 69
Db 1 CGAACGGACGTTTAAAGTGAGAAAAGAAACCGGTAAATCAGAGATCCCAAGCAAGCGCGT 60
Qy 70 CGGTGCATGATAGCGAAGAAAAAGCTATCCGTTTCAGTTAACTACTATTACCAAGATTGA 129
Db 61 CGGTGCATGATAGCGAAGAAAAAGCTATCCGTTTCAGTTAACTACTATTACCAAGATTGA 120
Qy 130 ATTTGCCCATCGGCAAAATTACTAAAATACATAAGTGCACACTGTCACCTGTGTGTGT 189
Db 121 ATTTGCCCATCGGCAAAATTACTAAAATACATAAGTGCACACTGTCACCTGTGTGTGT 180
Qy 190 GTTTTTTTTTTTTTTTTGGTTTTCGCTGTGCGCTTTTATCGCAAAACAAGACTGATAAAC 249
Db 181 GTTTTTTTTTTTTTTTTGGTTTTCGCTGTGCGCTTTTATCGCAAAACAAGACTGATAAAC 240
Qy 250 TAGAAAATATCTTGAGAAACTTGTTTTCGCGCTTTTCTTTTCTAAATGGCCGATCGCGGA 309
```

```
Db      241 TAGAAAATATCTTGAGAAACTTGTTTTCGCGCTTTTCTTTTGTCTAAATGGCCGATCGCGGA 300
Qy      310 AGAGAAAACAAGCAGTAGACAAAACAAGTGTGTAACAATCTGAAAAGGGGACCACATC 369
Db      301 AGAGAAAACAAGCAGTAGACAAAACAAGTGTGTAACAATCTGAAAAGGGGACCACATC 360
Qy      370 AGCAGCCCGAGGGGTTTATCTATATAGATGTCGCAGCTTATCATCTCATGCTGCTGTGA 429
Db      361 AGCAGCCCGAGGGGTTTATCTATATAGATGTCGCAGCTTATCATCTCATGCTGCTGTGA 420
Qy      430 GGTGTGTTCTGTGCTCGTGTAGTATCTTAAATACATAGAGTGTTCATATAAAGTCGG 489
Db      421 GGTGTGTTCTGTGCTCGTGTAGTATCTTAAATACATAGAGTGTGTTCATATAAAGTCGG 480
Qy      490 ACAAAGCTCGATTGGAACACAGCTGTCGAGTGCCCTTCAGTGGGTGGGCAAGATCGTCAATC 549
Db      481 ACAAAGCTCGATTGGAACACAGCTGTCGAGTGCCCTTCAGTGGGTGGGCAAGATCGTCAAT 540
Qy      550 ATCATCATCGTCGTCATTATCAACAGAAATCAGCATCAGCATCTCGAGGCCCGGTTGCTC 609
Db      541 ATCATCATCGGCGTCATTATCAACAGAAATCAGCATCAGCATTTTGAGGCCCGGTTGCTC 600
Qy      610 TAAG 613
Db      601 TAAG 604

RESULT 7
LOCUS   BI620499                565 bp mRNA linear EST 07-SEP-2001
DEFINITION RH51659.5prime RH Drosophila melanogaster normalized Head pFLC-1
Drosophila melanogaster cDNA clone RH51659 5, mRNA sequence.
ACCESSION BI620499
VERSION   BI620499.1 GI:15516024
KEYWORDS EST.
SOURCE   Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 565)
Stapleton,M., Brokstein,P., Hong,L., Tyler,D., Berman,B.,
Carlson,J., Champe,M., Chavez,C., Dorsett,V., Farfan,D., Frise,E.,
George,R., Gonzalez,M., Guarin,H., Harris,N., Li,P., Liao,G.,
Miera,S., Mungall,C.J., Munoo,J., Pacleb,J., Paragas,V., Park,S.,
Phouanavong,S., Wan,K., Yu,C., Lewis,S.E., Celniker,S. and
Rubin,G.M.
BDGP/HHMI RH Drosophila EST Project
Unpublished (2001)
Contact: Stapleton, M.
BDGP

Lawrence Berkeley National Lab
One Cyclotron Rd, Berkeley, CA 94720, USA
Fax: 510 486 6798
Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
hit genomic AE003831: arm:2R [4876906,5137815]
estimated-cyto:46B13-46E4: 08/22/2001
Plate: RH.516 row: E column: 11
High quality sequence stop: 442.
Location/Qualifiers
1..565
/organism="Drosophila melanogaster"
/mol_type="mRNA"
/db_xref="taxon:7227"
/clone="RH51659"
/sex="male and female"
/dev_stage="Adult"
/lab_host="DHS-alpha Tona"
/clone_lib="RH Drosophila melanogaster normalized Head
pFLC-1"
/notes="Organ: head; Vector: pFLC1; Site 1: XhoI; Site 2:
BamHI; Library was kindly generated by Piero Carninci at
```

the RIKEN. The library was normalized and excised using Cre recombinase. Plasmid cDNA library."

ORIGIN

Query Match 25.4%; Score 545.4; DB 12; Length 565;  
Best Local Similarity 99.8%; Pred. No. 6.8e-59;  
Matches 546; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 CGAGCGGACCGACGCTTTAAAGTGAGAAAAGAAACCGGTAAATCAGAGATCCCAAGCAAG 64  
Db |||  
19 CGACCGGACCGACGCTTTAAAGTGAGAAAAGAAACCGGTAAATCAGAGATCCCAAGCAAG 78  
QY 65 CGCGTGCCTGATGATAGCAGAAAGAAAAGCTATCCGTTTCAGTTAACTACTTACCAAG 124  
Db |||  
79 CGCGTGCCTGATGATAGCAGAAAGAAAAGCTATCCGTTTCAGTTAACTACTTACCAAG 138  
QY 125 ATTGAATTTCCCATCGGCGCAAAATTTACTTAAATAATACATAAGTGAACCTCGTCCACTGTGT 184  
Db |||  
139 ATTGAATTTCCCATCGGCGCAAAATTTACTTAAATAATACATAAGTGAACCTCGTCCACTGTGT 198  
QY 185 GTTGTGTTTTTTTTTTTTTTTTTTTTTTCGTTTCGCTGTGCTTTATCGCAACCAAGAACTGAT 244  
Db |||  
199 GTTGTGTTTTTTTTTTTTTTTTTTTTTTCGTTTCGCTGTGCTTTATCGCAACCAAGAACTGAT 258  
QY 245 AAAACTAGAAAATATCTTGAGAAACTGTGTTTCGCGCTTTTCTTTTGTCTAATTCGCGATC 304  
Db |||  
259 AAACTAGAAAATATCTTGAGAAACTGTGTTTCGCGCTTTTCTTTTGTCTAATTCGCGATC 318  
QY 305 GCGGAAGAGAAAAACAAGCAGTAGACAAAAACAAGTGTGTAATAACAATCTGAAAAGGGCA 364  
Db |||  
319 GCGGAAGAGAAAAACAAGCAGTAGACAAAAACAAGTGTGTAATAACAATCTGAAAAGGGCA 378  
QY 365 CCATCAGCAGCCGAGGGTTTATCTATATAGATGTCGACGTTATCATCTCATGCTGTC 424  
Db |||  
379 CCATCAGCAGCCGAGGGTTTATCTATATAGATGTCGACGTTATCATCTCATGCTGTC 438  
QY 425 TGTGAGGTTGTTCTGTGTCGTGTAGTATCTTAAATACATAGAGTGTTCATATAA 484  
Db |||  
439 TGTGAGGTTGTTCTGTGTCGTGTAGTATCTTAAATACATAGAGTGTTCATATAA 498  
QY 485 GTGCGACAAAGCTCGATTGAAAACAGCTGTCGAGTGCCTTGAGTGGGTGGGCAAGATCG 544  
Db |||  
499 GTGCGACAAAGCTCGATTGAAAACAGCTGTCGAGTGCCTTGAGTGGGTGGGCAAGATCG 558  
QY 545 TCATCAT 551  
Db 559 TCATCAT 565

RESULT 8  
CKI35985/c  
LOCUS  
DEFINITION  
SD18286.3prime SD Drosophila melanogaster Schneider L2 cell culture  
pOT2 Drosophila melanogaster cDNA clone SD18286 3, mRNA sequence.  
CKI35985  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Drosophila melanogaster (fruit fly)  
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
Ephydroidea; Drosophilidae; Drosophila.  
1 (bases 1 to 508)  
Harvey, D., Brokstein, P., Hong, L., Evans-Holm, M., Su, C., Tsang, G.,  
Lewis, S. and Rubin, G.M.  
BDGP/HMI Drosophila EST Project  
Unpublished (2001)  
Other ESTs: SD18286.5prime  
Contact: Stapleton, M.  
BDGP  
Lawrence Berkeley National Lab  
One Cyclotron Rd, Berkeley, CA 94720, USA  
Fax: 510 486 6798

Email: <http://www.fruitfly.org/EST>, [est@fruitfly.berkeley.edu](mailto:est@fruitfly.berkeley.edu)  
Based upon the presence of a XhoI site followed by a run of 14 or more T residues at the beginning of the sequence, this clone was polyadenylated. The resulting Poly-T sequence has been removed. hit genomic AE003831: arm:2R [4876906..5137815]  
estimated-cyto:46B13-46B4: 02/08/2002  
Plate: SD.182 row: H column: 2  
High quality sequence stop: 456.  
Location/Qualifiers  
1..508  
/organism="Drosophila melanogaster"  
/mol\_type="mRNA"  
/db\_xref="taxon:7227"  
/clone="SD18286"  
/lab\_host="DHS-alpha"  
/clone\_lib="SD Drosophila melanogaster Schneider L2 cell culture pOT2"  
/note="Vector: pOT2; Site 1: EcoRI; Site 2: XhoI; Sized fractionated cDNAs were directly ligated into pOT2. Plasmid cDNA library."

ORIGIN

Query Match 23.3%; Score 500; DB 14; Length 508;  
Best Local Similarity 99.0%; Pred. No. 3.3e-53;  
Matches 503; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1603 TATTACGTATAGCCCGAGATATGCTTAAACAATCTGCACGACGAGATTTATCGTC 1662  
Db |||  
508 TATTACGTATAGCCCGAGATATGCTTAAACAATCTGCACGACGAGATTTATCGTC 449  
QY 1663 TTTCAAGGAGACATCCATTCTCGAGTGCTTGAACACGGTCCCAACCAATCCACAT 1722  
Db |||  
448 TTTCAAGGAGACATCCATTCTCGAGTGCTTGAACACGGTCCCAACCAATCCACAT 389  
QY 1723 AAGGTGCACACCTGCCACACGAGTGGTCTGTATCCACCTGGAAACGAGAGGATCCAT 1782  
Db |||  
388 AAGGTGCACACCTGCCACACGAGTGGTCTGTATCCACCTGGAAACGAGAGGATCCAT 329  
QY 1783 CTGAAGGACATTCACAAAGATCGAATCGAATGCTTTCGGGAGGGAAACACCGAAGCTAC 1842  
Db |||  
328 CTGAAGGACATTCACAAAGATCGAATGCTTTCGGGAGGGAAACACCGAAGCTAC 269  
QY 1843 TTTTGCATCTTCAAGGTGTAAATTCGAGAGATTATCCCGGTGCAAGATGAATACCAAG 1902  
Db |||  
268 TTTTGCATCTTCAAGGTGTAAATTCGAGAGATTATCCCGGTGCAAGATGAATACCAAG 209  
QY 1903 TTTAAGCTTTTGTCCCGCACTGCTCGTGAATGCGATTTCATCGCCAGCGTAATCCATT 1962  
Db |||  
208 TTTAAGCTTTTGTCCCGCACTGCTCGTGAATGCGATTTCATCGCCAGCGTAATCCATT 149  
QY 1963 AGTTCTGTAGTACCTAGTCTTAGTCACTCAAAACCTAATCTCAATCGGAATCGTGATACT 2022  
Db |||  
148 AGTTCTGTAGTACCTAGTCTTAGTCACTCAAAACCTAATCTCAATCGGAATCGTGATACT 89  
QY 2023 GCATTAGTCAGAAAGACGGAGGAAATCATTTATTTTGTATATACTCGTTTCGACTCTAA 2082  
Db |||  
88 GCATTAGTCAGAAAGACGGAGGAAATCATTTATTTTGTATATACTCGTTTCGACTCTAA 29  
QY 2083 AAAGTGAATAAAATATATATAGTCTATT 2110  
Db |||  
28 AAAGTGAATAAAATATATATAGTCTATT 1

RESULT 9  
CKI35525/c  
LOCUS  
DEFINITION  
RH51659.3prime RH Drosophila melanogaster normalized Head pR1c-1  
Drosophila melanogaster cDNA clone RH51659 3 similar to CG12919:  
FBan0012919 GO:[] located on: 2R 46E1-46E1:: 08/05/2002, mRNA  
sequence.  
CKI35525  
ACCESSION  
VERSION  
KEYWORDS  
EST.

```
SOURCE
ORGANISM
Drosophila melanogaster (fruit fly)
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 532)
REFERENCE
AUTHORS
Stapleton,M., Brokstein,P., Hong,L., Tyler,D., Berman,B.,
Carlson,J., Chamez,C., Chavez,C., Dorsett,V., Farfan,D., Frise,E.,
George,R., Gonzalez,M., Guarin,H., Harris,N., Li,P., Liao,G.,
Miera,S., Mungali,C.J., Nunoo,J., Pacleb,J., Paragas,V., Park,S.,
Phouanavong,S., Wan,K., Yu,C., Lewis,S.E., Celniker,S. and
Rubin,G.M.
BDGP/HHMI RH Drosophila EST Project
Unpublished (2001)
Other ESTs: RH51659-5prime
Contact: Stapleton, M.
BDGP
Lawrence Berkeley National Lab
One Cyclotron Rd, Berkeley, CA 94720, USA
Fax: 510 486 6798
Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
Based upon the presence of a XhoI site followed by a run of 14 or
more T residues at the beginning of the sequence, this clone was
polyadenylated. The resulting Poly-T sequence has been removed. hit
genomic AE003831: arm:2R [4876906,5137815]
estimated-cyto:46B13-46E4: 08/05/2002
Plate: RH.516 row: E column: 11
High quality sequence stop: 382.
FEATURES
source
1..532
/organism="Drosophila melanogaster"
/mol_type="mRNA"
/db_xref="taxon:7227"
/clone="RH51659"
/sex="male and female"
/dev_stage="Adult"
/lab_host="DHS-alpha Tona"
/clone_lib="RH Drosophila melanogaster normalized Head
pFle-1"
/note="Organ: head; Vector: pFle1; Site:1: XhoI; Site:2:
BamHI; Library was kindly generated by Piero Carninci at
the RIKEN. The library was normalized and excised using
Cre recombinase. Plasmid cDNA library."
ORIGIN
Query Match 23.3%; Score 500; DB 14; Length 532;
Best Local Similarity 99.0%; Pred. No. 3.2e-53;
Matches 503; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 1604 ATTAGTATAGCCCGAGATATGCTACAACTCGCAGCAGCAACGGATTTATCGTCT 1663
DB 532 ATTAGTATAGCCCGAGATATGCTACAACTCGCAGCAGCAACGGATTTATCGTCT 473
QY 1664 TTCAAGGAGACATCCATCTCTGCGAGTGCTTGAACACCGTGCCCAACCAATGCCACATA 1723
DB 472 TTCAAGGAGACATCCATCTCTGCGAGTGCTTGAACACCGTGCCCAACCAATGCCACATA 413
QY 1724 AGGTGCACCTGCCACACGAGTGCTGTGATCCACCTGGAACGAGAGGATCCATC 1783
DB 412 AGGTGCACCTGCCACACGAGTGCTGTGATCCACCTGGAACGAGAGGATCCATC 353
QY 1784 TGAAGGACATTCACAAAGATCGCAATGCGAGTTCTCGGGAGGGAACCAACGAGCTACT 1843
DB 352 TGAAGGACATTCACAAAGATCGCAATGCGAGTTCTCGGGAGGGAACCAACGAGCTACT 293
QY 1844 TTGGCATCTTCAAGGTGTAATTGGAGAGATTATCCCGGTGAGAAAGATGAATACCAAGT 1903
DB 292 TTGGCATCTTCAAGGTGTAATTGGAGAGATTATCCCGGTGAGAAAGATGAATACCAAGT 233
QY 1904 TTAGCTTTTGTCCCGGACTGCTCGTGAATGCGATTTCATCGCCAGCGTGAATCCATT 1963
DB 232 TTAAGCTTTTGTCCCGGACTGCTCGTGAATGCGATTTCATCGCCAGCGTGAATCCATT 173
```

```
QY 1964 GTTCGTAGTACTAGTCTTAGTCACTCAACACCTAATCTCAATCGGAATCGTGACTACTG 2023
DB 172 GTTCGTAGTACTAGTCTTAGTCACTCAACACCTAATCTCAATCGGAATCGTGACTACTG 113
QY 2024 CATTAGTCAGAACGAGCGAGGAAATCATATTATTTGTATATATCTCGTTGACTCTTAA 2083
DB 112 CATTAGTCAGAACGAGCGAGGAAATCATATTATTTGTATATATCTCGTTGACTCTTAA 53
QY 2084 AAGTGAATATAAATATATATGTAGCTATT 2111
DB 52 AAGTGAATATAAATATATATGTAGCTATT 25
RESULT 10
BI636564
LOCUS
DEFINITION
SD18286.Sprime SD Drosophila melanogaster Schneider L2 cell culture
pot2 Drosophila melanogaster cDNA clone SD18286 5 similar to
CG12919: FBan0012919 GO:[] located on: 2R 46E1-46E1; 05/19/2001,
mRNA sequence.
ACCESSION
BI636564
VERSION
BI636564.1 GI:15530774
KEYWORDS
EST.
ORGANISM
Drosophila melanogaster (fruit fly)
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 495)
REFERENCE
AUTHORS
Harvey,D., Brokstein,P., Hong,L., Evans-Holm,M., Su,C., Tsang,G.,
Lewis,S. and Rubin,G.M.
BDGP/HHMI Drosophila EST Project
Unpublished (2001)
Contact: Stapleton, M.
BDGP
Lawrence Berkeley National Lab
One Cyclotron Rd, Berkeley, CA 94720, USA
Fax: 510 486 6798
Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
hit genomic AE003831: arm:2R [4876906,5137815]
estimated-cyto:46B13-46E4: 05/19/2001
Plate: SD.182 row: H column: 2
High quality sequence stop: 491.
FEATURES
source
1..495
/organism="Drosophila melanogaster"
/mol_type="mRNA"
/db_xref="taxon:7227"
/clone="SD18286"
/lab_host="DHS-alpha"
/clone_lib="SD Drosophila melanogaster Schneider L2 cell
culture pot2"
/note="vector: pot2; Site:1: EcoRI; Site:2: XhoI; Sized
fractionated cDNAs were directly ligated into pot2.
Plasmid cDNA library."
ORIGIN
Query Match 22.2%; Score 477.4; DB 12; Length 495;
Best Local Similarity 99.6%; Pred. No. 2.2e-50;
Matches 489; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
QY 931 TACGAGAATGCCCTCATCGACTATCCAAAAAGGTGGATGGCCCTCAGGATGAGGAGAC 990
DB 6 TACGAGAATGCCCTCATCGACTATCC-CAAAAGGTGGATGGCCCTCAGGATGAGGAGAC 64
QY 991 GACGACGATGCGATGCTCTGGATTCCATTTCGCGACGACGAGGACGACGCTTAGCTAT 1050
DB 65 GACGACGATGCGATGCTCTGGATTCCATTTCGCGACGACGAGGACGACGCTTAGCTAT 124
QY 1051 AGCTCTGTGATGATGTTGGCGCAGACTACGAGACTATGTTAAATAAATCTC 1110
DB 125 AGCTCTGTGATGATGTTGGCGCAGACTACGAGACTATGTTAAATAAATCTC 184
```



/clone\_lib="Glossina morsitans morsitans adult infected gut"  
/note="country: Zimbabwe; EST from adult gut infected with T.brucei"

```
ORIGIN
Query Match      6.0%; Score 129.8; DB 13; Length 612;
Best Local Similarity 64.5%; Pred. No. 3.9e-07;
Matches 194; Conservative 0; Mismatches 107; Indels 0; Gaps 0;

Qy 1571 ATGGCGCTTACCGGTGACCAATACAGCCCTATTAGCTATACGCCAGATATGCTACA 1630
Db 607 ATGGTGTCTTGAATTCGGAATCGGCTTTTATTATGTTTATGCTCAAGTATGTTATA 548

Qy 1631 ACAACTCCACGACGACAGAACGGATTATCGTCTTTCAGAGGAGACACTCCATTCCTGCAGT 1690
Db 547 ATACACGACGATCATTAATGGTTTGTGTTATTCATGGTCAAGCACCTTTTACAAAT 488

Qy 1691 GCTTGAACACCGGTGCCACCAACATGCCACATAAGGTGCACACCTGCCACACGAGTGGTC 1750
Db 487 GCTTCAAACTGTACCCACTAACATGTCACTGAAATCCACACTTGTCTACGAGCGGT 428

Qy 1751 TGATCCACCTGGACGAAACGAGAGGATCCATCTGAAGGACATTCACAAACGATCGCAATG 1810
Db 427 TAATTTATTTAAAAACACATGAAATAATTTACATTTAAGAGATTTTACACAGTGACCGCAATG 368

Qy 1811 CAGTTCTCGGGAGGGAACCAACGAGCTACTTTGGCATCTTCAAGGTGTAAATTTGAG 1870
Db 367 CGGTCTGAAGGATTCAAACAATAGAAGTTACTTTGGTTTAAATGAATAATCTGAAAGAAA 308

Qy 1871 A 1871
Db 307 A 307

RESULT 13
CNS017WQ      1101 bp DNA linear GSS 26-JUL-1999
LOCUS      Drosophila melanogaster genome survey sequence T7 end of BAC
DEFINITION      BACN37D20 of DrosBAC library from Drosophila melanogaster (fruit fly), genomic survey sequence.
ACCESSION      AL108596.1 GI:5628900
VERSION      GSS.
KEYWORDS      Drosophila melanogaster (fruit fly)
SOURCE      Drosophila melanogaster
ORGANISM      Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE      1 (bases 1 to 1101)
AUTHORS      Genoscope.
TITLE      Direct Submission
JOURNAL      Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)
COMMENT      Determination of this BAC-end sequence was carried out as part of a collaboration with the European Drosophila Genome Project (EDGP) - http://www.edgp.ebi.ac.uk -- This Drosophila melanogaster BAC library (Dros BAC) was made by Alain Billaud at CEPH (Centre d'Etude du Polymorphisme Humain) with funding provided by a MRC project grant. The DNA was prepared from embryos by Alain Bucheton and Genevieve Payan. It has been constructed in the vector pBelobAC11.
FEATURES      Location/Qualifiers
source      1..1101
             /organism="Drosophila melanogaster"
             /mol_type="genomic DNA"
             /db_xref="taxon:7227"
             /clone="BACN37D20"
             /clone_lib="DrosBAC"
             /plasmid="pBelobAC11"
             /note="end : T7"

ORIGIN
Query Match      3.1%; Score 65.6; DB 12; Length 468;
Best Local Similarity 48.9%; Pred. No. 47;
Matches 176; Conservative 0; Mismatches 184; Indels 0; Gaps 0;

Qy 859 CGAGTCTCGATATATCTCCAGCAGCGTTTGGGCATAAATCTCTGGCAGATTCGACGAG 918
Db 21 CGTGACGATGATGATGACGACGATGATGACGACGATGATGATGATGATGATGATGATGAC 80
```

```
Query Match      5.4%; Score 115.6; DB 29; Length 1101;
Best Local Similarity 72.0%; Pred. No. 1.5e-05;
Matches 131; Conservative 13; Mismatches 38; Indels 0; Gaps 0;

Qy 804 GATTCGCAGACAAACGCGTGTATCGCATCTGGACAAGAGCTGAAGACGCTGAAGCGAGT 863
Db 261 GAWMTGGCAGAAAMRCGTGTATCGCATCTGGACAAAGAACTGAAAAACCTGAAVCART 320

Qy 864 CGTCGATAATCTCCAGCAGCGTTTGGGCATAAACTATCTGCAGCAGTTTCGACGAGTTCCA 923
Db 321 CGTCRATAATCTCCAGCAACGTTTGGGCATAAACTATCTGCAGCAATACGACGATTTCCA 380

Qy 924 AAAGAGGTACGAGATGCCCTCATCGCATATCAAAAAAGGTGGATGGCCTCACGGATGA 983
Db 381 AAAGAGGTACGAGATGCCCTCATCGCATATCAAAAAAGGTGGATGGCCTCACGGATGA 440

Qy 984 GG 985
Db 441 GG 442

RESULT 14
BG553005      468 bp mRNA linear EST 09-APR-2001
LOCUS      dab82e09.y1 NICHD_XGC_Emb4 Xenopus laevis cDNA clone IMAGE:4203736
DEFINITION      5', mRNA sequence.
ACCESSION      BG553005
VERSION      BG553005.1 GI:13564785
KEYWORDS      EST.
SOURCE      Xenopus laevis (African clawed frog)
ORGANISM      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae; Xenopodinae; Xenopus.
REFERENCE      1 (bases 1 to 468)
AUTHORS      NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE      National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL      Unpublished (1997)
COMMENT      Contact: Robert Strausberg, Ph.D.
             Email: cgapbs-rc@mail.nih.gov
             Tissue Procurement: Martha Rebert, Steven L. Klein, Ph.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: xenopus clones from this library are available through the I.M.A.G.E. Consortium/LLNL at: info@image.llnl.gov
             Seq primer: -40RP from Gibco
             High quality sequence stop: 404.
FEATURES      Location/Qualifiers
source      1..468
             /organism="Xenopus laevis"
             /mol_type="mRNA"
             /db_xref="taxon:8355"
             /clone="IMAGE:4203736"
             /dev_stage="embryo stage 31-32"
             /lab_hosts="DH10B (phage-resistant)"
             /clone_lib="NICHD_XGC_Emb4"
             /notes="Organ: whole embryo; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 2.1 kb. Constructed by Life Technologies. Note: This is a Xenopus Gene Collection (XGC) library."

ORIGIN
Query Match      3.1%; Score 65.6; DB 12; Length 468;
Best Local Similarity 48.9%; Pred. No. 47;
Matches 176; Conservative 0; Mismatches 184; Indels 0; Gaps 0;

Qy 859 CGAGTCTCGATATATCTCCAGCAGCGTTTGGGCATAAATCTCTGGCAGATTCGACGAG 918
Db 21 CGTGACGATGATGATGACGACGATGATGACGACGATGATGATGATGATGATGATGATGAC 80
```





GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 13, 2004, 10:21:32 ; Search time 57 Seconds  
(without alignments)  
2027.402 Million cell updates/sec

Title: US-09-813-329-6

Perfect score: 2162

Sequence: 1 MTAETLKPFTTPTTSANDGF.....DRNAVLREGNRSYFGIFKV 409

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_29Jan04.\*  
1: geneseqp1980s.\*  
2: geneseqp1990s.\*  
3: geneseqp2000s.\*  
4: geneseqp2001s.\*  
5: geneseqp2002s.\*  
6: geneseqp2003as.\*  
7: geneseqp2003bs.\*  
8: geneseqp2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	2162	100.0	409	5	Aau77718	Drosophil
2	2082.5	96.3	406	5	Aau77717	Aau77717 Drosophil
3	2045	94.6	409	5	Aau77716	Aau77716 Drosophil
4	1656	76.6	325	4	Abb67553	Abb67553 Drosophil
5	149	6.9	27	5	Aau88008	Aau88008 Tumour ne
6	149	6.9	27	5	Aau88018	Aau88018 Tumour ne
7	141.5	6.5	330	5	Abg32871	Abg32871 Mouse Zis
8	141	6.5	428	5	Abg93245	Abg93245 C. albica
9	138.5	6.4	1192	3	Aab18165	Aab18165 Plasmodiu
10	138.5	6.4	1883	4	Abb60038	Abb60038 Drosophil
11	137	6.3	330	3	Aay70227	Aay70227 Human RNA
12	137	6.3	532	4	Abb67173	Abb67173 Drosophil
13	137	6.3	2175	4	Abb65698	Abb65698 Drosophil
14	135	6.2	330	4	Aau21828	Aau21828 Novel hum
15	135	6.2	330	4	Aau86946	Aau86946 Human DNA
16	135	6.2	330	7	ADC46469	ADC46469 Human neo
17	134	6.2	466	6	ABR52760	ABR52760 Protein s
18	133.5	6.2	3583	4	ABR64814	ABR64814 Drosophil
19	133	6.2	234	4	AAB94508	AAB94508 Human pro
20	133	6.2	2500	3	AAB18272	AAB18272 Plasmodiu
21	132	6.1	391	5	ABP53104	ABP53104 Human ect
22	132	6.1	391	5	ABR09082	ABR09082 Human EDA
23	132	6.1	391	6	ABR42324	ABR42324 Human EDA
24	132	6.1	391	7	ABO01928	ABO01928 Human EDA
25	132	6.1	391	7	ADC35224	ADC35224 Human TNF

ALIGNMENTS

RESULT 1

AAU77718

ID AAU77718 standard; protein; 409 AA.

XX AAU77718;

AC AAU77718;

DT 05-JUN-2002 (first entry)

DE Drosophila melanogaster tumour necrosis factor variant 2 (TNFv2).

KW Tumour necrosis factor; TNF; pesticide; crop-damaging insect;  
KW farm animal insect; epithelial morphogenesis; cell-matrix adhesion;  
KW X-linked anhidrotic (hypohidrotic) ectodermal dysplasia;  
KW X-linked anhidrotic (hypohidrotic) ectodermal dysplasia-like disorder;  
KW sparse hair; sweat gland aberration; endotoxin shock; inflammation;  
KW haemorrhagic necrosis of tumour; cytotoxicity; TNFv2;  
KW obesity-linked insulin resistance.

XX Drosophila melanogaster.

XX Key Location/Qualifiers  
FT Peptide 1..52  
FT Protein 53..409  
/label= Signal\_peptide  
/label= mature\_TNF  
/note= "Tumour necrosis factor. Specifically claimed in claim 36"

US2002012968-A1.

31-JAN-2002.

20-MAR-2001; 2001US-00813329.

21-MAR-2000; 2000US-0190816P.

(CARR/) CARROLL P M.

(CHEN/) CHEN J.

(RAMA/) RAMANATHAN C S.

(XIAO/) XIAO H.

(GUAN/) GUAN B.

(BOWE/) BOWEN M A.

Carroll PM, Chen J, Ramanathan CS, Xiao H, Guan B, Bowen MA;

WPI; 2002-195121/25.

N-PSDB; ABK11680.

XX



Db 121 DGGDGLSDIADDEDDDDVSYSSVDDYGDYEDYDMLNKLNNNAHTGTTPTSETTAEGEGET 180  
Qy 181 DSASSASNDNDVDDFTSYNAHKKQERKRSIADVRNEEQNIQGNHTELOEKSSNEATS 240  
Db 181 DSASSASNDNDVDDFTSYNAHKKQERKRSIADVRNEEQNIQGNHTELOEKSSNEATS 240  
Qy 241 KESPAPLHRRMRHSHRHLVRK-----ARSEDSRPAAHFHLSSRRRHQSGMGYHGDM 294  
Db 241 KE-----RMSHRHLLVRKGESLRSARSEDSRPAAHFHLSSRRRHQSGMGYHGDM 291  
Qy 295 YIGNDNERNVSQCHPQTRDGLVTNTGLYYVYQAICYNNSHDQNGFVFGDTPFLQCL 354  
Db 292 YIGNDNERNVSQGHFQTRDGLVTNTGLYYVYQAICYNNSHDQNGFVFGDTPFLQCL 351  
Qy 355 NTVPTNMPKHVHTCHTSLIHLERNERHLKDIHNDNRNAVLRGNRRSYFGIFKV 409  
Db 352 NTVPTNMPKHVHTCHTSLIHLERNERHLKDIHNDNRNAVLRGNRRSYFGIFKV 406

RESULT 3  
AAU77716  
ID AAU77716 standard; protein; 409 AA.  
XX  
AC AAU77716;  
XX  
DT 05-JUN-2002 (first entry)  
XX  
DE Drosophila melanogaster tumour necrosis factor (TNF).  
XX  
KW Tumour necrosis factor; TNF; pesticide; crop-damaging insect;  
KW farm animal insect; epithelial morphogenesis; cell-matrix adhesion;  
KW X-linked anhidrotic (hypohidrotic) ectodermal dysplasia;  
KW X-linked anhidrotic (hypohidrotic) ectodermal dysplasia-like disorder;  
KW sparse hair; sweat gland aberration; endotoxic shock; inflammation;  
KW haemorrhagic necrosis of tumour; cytotoxicity;  
KW obesity-linked insulin resistance.  
XX  
OS Drosophila melanogaster.  
XX  
FH Key Location/Qualifiers  
FT Peptide 1..52  
FT Protein /label= Signal\_peptide  
FT /label= mature TNF  
FT /note= "Tumour necrosis factor. Specifically claimed in  
FT claim 34"

US2002012968-A1.  
31-JAN-2002.  
20-MAR-2001; 2001US-00813329.  
21-MAR-2000; 2000US-0190816P.  
(CARR/) CARROLL P M.  
PA (CHEN/) CHEN J.  
PA (RAMA/) RAMANATHAN C S.  
PA (XIAO/) XIAO H.  
PA (GUAN/) GUAN B.  
PA (BOWE/) BOWEN M A.  
XX  
PI Carroll PM, Chen J, Ramanathan CS, Xiao H, Guan B, Bowen MA;  
XX WPI; 2002-195121/25.  
DR N-PSDB; ABK11678.  
XX  
PT New Drosophila tumor necrosis factor molecule, useful in controlling  
PT agriculturally important pests, e.g. comprises modifying the growth,  
PT feeding or reproduction of crop-damaging insects or insects of farm  
PT animals.  
XX

PS Claim 6; Fig 1A-C; 119pp; English.  
XX  
CC The invention describes an isolated tumour necrosis factor polypeptide  
CC (TNF). The polypeptide and polynucleotide are useful in controlling  
CC agriculturally important pests, particularly by modifying the growth,  
CC feeding and/or reproduction of crop-damaging insects or insects of farm  
CC animals. The polypeptide and polynucleotide are useful for modulating  
CC epithelial morphogenesis, cell-matrix adhesion in flies and mammals. Thus  
CC the polypeptide and polynucleotide may be useful for treating,  
CC ameliorating or preventing X-linked anhidrotic (hypohidrotic) ectodermal  
CC dysplasia and X-linked anhidrotic (hypohidrotic) ectodermal dysplasia-  
CC like disorders, e.g. sparse hair, abnormal or missing teeth or sweat  
CC gland aberrations in animals (e.g. insects and potentially humans),  
CC endotoxic shock, inflammation, haemorrhagic necrosis of tumours,  
CC cytotoxicity and obesity-linked insulin resistance, all of which involve  
CC TNF molecules. This is the amino acid sequence of the Drosophila  
CC melanogaster tumour necrosis factor protein, described in the invention  
XX  
SQ Sequence 409 AA;  
Query Match 94.6%; Score 2045; DB 5; Length 409;  
Best Local Similarity 96.3%; Pred. No. 1.5e-171;  
Matches 394; Conservative 2; Mismatches 13; Indels 0; Gaps 0;  
Qy 1 MTAE TLKPFITPTTSANDDGFPAKATSTATAQRTRQLPLVLGFIGLVVAILALTIWQ 60  
Db 1 MTAE TLKPFITPTTSANDDGFPAKATSTATAQRTRQLPLVLGFIGLVVAILALTIWQ 60  
Qy 61 TTRVSHLDKELKSLKRVVDNLQORLGINYLDEFDFQKEYENALIDYPPKVDGLTDEDD 120  
Db 61 TTRVSHLDKELKSLKRVVDNLQORLGINYLDEFDFQKEYENALIDYPPKVDGLTDEDD 120  
Qy 121 DGGDGLSDIADDEDDDDVSYSSVDDVGDYEDYDMLNKLNNNAHTGTTPTSETTAEGEGET 180  
Db 121 DGGDGLSDIADDEDDDDVSYSSVDDVGDYEDYDMLNKLNNNAHTGTTPTSETTAEGEGET 180  
Qy 181 DSASSASNDNDVDDFTSYNAHKKQERKRSIADVRNEEQNIQGNHTELOEKSSNEATS 240  
Db 181 DSASSASNDNDVDDFTSYNAHKKQERKRSIADVRNEEQNIQGNHTELOEKSSNEATS 240  
Qy 241 KESPAPLHRRMRHSHRHLVRKARSEDSRPAAHFHLSSRRRHQSGMGYHGDMYICNDN 300  
Db 241 KESPAALHLRRMRHSHRHLVRKARSEDSRPAAHFHLSSRRRHQSGMGYHGMYIENDR 300  
Qy 301 ERNSYQGHFOTRGVLTNTGLYYVYQAICYNNSHDQNGFVFGDTPFLQCLNTVPTN 360  
Db 301 ERCSYQGHFOTRGVLTNTNAGLYYYVYQAIWGYNSHDQNGFVFGDTPFLQCLNTVPTN 360  
Qy 361 MPHKVHTCHTSGLIHLERNERHLKDIHNDNRNAVLRGNRRSYFGIFKV 409  
Db 361 MPHKVHTCHTSGLIHLERNERHLKDIHNDNRNAVLRGNRRSYFGIFKV 409  
RESULT 4  
ABB67553  
ID ABB67553 standard; protein; 325 AA.  
XX  
AC ABB67553;  
XX  
DT 26-MAR-2002 (first entry)  
XX  
DE Drosophila melanogaster polypeptide SEQ ID NO 29451.  
XX  
KW Drosophila; developmental biology; cell signalling; insecticide;  
KW pharmaceutical.  
XX  
OS Drosophila melanogaster.  
XX  
PN WO200171042-A2.  
XX  
PD 27-SEP-2001.  
XX  
PF 23-MAR-2001; 2001WO-US009231.





QY 111 V-----DGLTDEDDDDGGLDLSIADDEDDVSVSSVDVGVADYEDYDMLKLN----- 161  
 Db 144 ILKEVEDKESEGEDEDELSKYKLDDEDEDDADLSKYNLDASEEDSNKKSRNR 203  
 QY 162 -----AHTGTTPTSETTAGEGETDGSASSANDNDNVFDDFTSYNAHKKQERKRSIA 214  
 Db 204 SKRSRSHRSRSPSSSRSRSRSRSSSSQSR---SHSGSRHSRSGSKRS-- 258  
 QY 215 DVNEEQNIQNHTELQKSNKATSKESAPLHRRHRRHRLVVRKARSDSR-PA 273  
 Db 259 -----SSRHRGSSSRPKRSYSSSSSS-----PERDKRSRSPSPSPAVKRRTRRSRPE 309  
 QY 274 AHFLSSRRRHGS 287  
 Db 310 RHRHSSSGSTHGS 323

RESULT 8  
 ID ABG93245 standard; protein; 428 AA.  
 XX AC ABG93245;  
 XX DT 21-NOV-2002 (first entry)  
 XX C. albicans BAX-associated protein fragment SEQ ID 448.  
 KW Bax; Bax-resistance; cytostatic; fungicide; immunosuppressive; virucide;  
 KW vasotropic; vaccine; gene therapy; proliferative disorder; cancer;  
 KW apoptosis; fungal; yeast; infection; autoimmune disease; ischaemia;  
 KW neurodegeneration; cell death.  
 XX Candida albicans.  
 XX WO200264766-A2.  
 XX FN 22-AUG-2002.  
 XX PD 21-DEC-2001; 2001WO-EP015398.  
 XX PF 22-DEC-2000; 2000EP-00870318.  
 XX PR 04-JAN-2001; 2001EP-00870002.  
 XX PR 09-JAN-2001; 2001EP-00870003.  
 XX (JANC ) JANSEN PHARM NV.  
 XX PA Contreras RH, Eberhardt I, Luyten WHML, Reekmans RJ;  
 XX PI WPI; 2002-667002/71.  
 XX DR N-PSDB; ABQ76511.

XX New isolated nucleic acid representing a synthetic BAX-gene, useful as  
 PT medicament for treating, preventing and/or alleviating yeast or fungal  
 PT infections or proliferative disorders, or for preventing apoptosis in  
 PT certain diseases.

XX Claim 36; Fig 2; 344pp; English.

XX This invention describes a novel nucleic acid representing a synthetic  
 CC Bax gene. The Bax gene of the invention is useful for identifying Bax-  
 CC resistant yeast or fungi, identifying, or obtaining and identifying  
 CC Candida spp. sequences that are differentially expressed in a pathway  
 CC eventually leading to programmed cell death or identifying inhibitors or  
 CC inhibitor sequences of Bax-induced cell death. The products of the  
 CC invention have cytostatic, fungicide; immunosuppressive, virucide and  
 CC vasotropic activity and can be used in vaccines or for gene therapy. The  
 CC isolated nucleic acids, polypeptides, pharmaceutical compositions,  
 CC antisense molecules and antibodies are useful as medicaments or in  
 CC preparing a medicament for treating, preventing and/or alleviating  
 CC diseases associated with yeast or fungi or proliferative disorders, such  
 CC as cancer, or for preventing apoptosis in certain diseases. The compounds  
 CC or polypeptides, or the genetically modified organism are useful for

CC preparing a medicament for modifying the endogenic flora of humans and  
 CC other mammals. The vaccine is useful for immunising against yeast or  
 CC fungal infections. Apoptosis-related diseases include autoimmune disease,  
 CC ischaemia, diseases related with viral infections or neurodegenerations.  
 CC This sequence represents a polypeptide associated with the Bax gene  
 CC described in the disclosure of the invention

XX Sequence 428 AA;

Query Match 6.5%; Score 141; DB 5; Length 428;  
 Best Local Similarity 20.8%; Pred. No. 0.0015;  
 Matches 76; Conservative 68; Mismatches 138; Indels 84; Gaps 16;

QY 36 QLIPLV-LGFIGLGLVAILALTIWQTTTVSHLDKEL-----KSLKRVVDNL 81  
 Db 66 QLLPIAKMSNTQDLVAYINDYVSRNEELSKLKALSKFLAGKELPKVKQLESIIDEV 125  
 QY 82 QQR----LGINYLDDEFQKEVEYALIDYPKK---VDGLTDEDDDDGGLDSTADDED 134  
 Db 126 ENQEKSKPRNSSDSESSSESTSDSSSDSSSDSSSDSSSDSSSDSSSDSSDSD 185  
 QY 135 DVVSYSSVD-DVGADYEDYDMLN-KLNNAHGTGTTPTSETTAGEGETDSASSANDNV 192  
 Db 186 SDDEEDKEDKAEKDKNGKSEDSENEKVEDNKTSSDSSSDSSSDSSSDSSSDSS 245  
 QY 193 FDFFTSYNAHKKQERKRSIADVNE---EQNIQNHTEL---QEKSSNEATSKESPAP 246  
 Db 246 DSDSSSDSSSDSSSDSSSDSSSDSSSDSSSDSSSDSSSDSSSDSSSDSSSESESE 305  
 QY 247 LHHRRMRHSR--HRLHLVRKARSE-----DSRPAA-----HFHLSRRR- 283  
 Db 306 PEDKKRKHDTDDIKEEPKPKFKNESESSASSSTSDSIPATPELPKGQKHFSDRSKV 365  
 QY 284 -----HOGSNMGYHGM-----YIGNDN--ERNYSQGHFQTRDGLVLT 318  
 Db 366 NFENSVLQDNTYKGAAGTWGEKASEKLLQVRGKDFTKNKNMKGYSYKG-----GSITL 419  
 QY 319 TNTGLY 324  
 Db 420 A-SGSY 424

RESULT 9

AA18165  
 ID AAB18165 standard; protein; 1192 AA.

XX AC AAB18165;

XX DT 07-NOV-2000 (first entry)

XX DE Plasmodium falciparum chromosome 2 related protein SEQ ID NO:22.

XX Plasmodium falciparum; chromosome 2; human malaria parasite; vaccine;  
 KW antimalarial; malaria; protozoacide; infection; insecticide.

XX OS Plasmodium falciparum.

XX FN WO200025728-A2.

XX PD 11-MAY-2000.

XX PF 05-NOV-1999; 99WO-US026796.

XX PR 05-NOV-1998; 98US-0107131P.

XX (HOFF/) HOFFMAN S.

PA (CARU/) CARUCCI D.

PA (GARD/) GARDNER M.

PA (VENT/) VENTER J C.

XX PI Hoffman S, Carucci D, Gardner M, Venter JC;  
 XX WPI; 2000-365347/31.

DR







XX ABB67173;  
XX AC  
XX DT  
XX DE  
XX DE  
XX KW  
XX KW  
XX OS  
XX PN  
XX PD  
XX PF  
XX PR  
XX PR  
XX PA  
XX PI  
XX DR  
XX DR  
XX PT  
XX PT  
XX PS  
XX PS  
XX CC  
XX CC  
XX CC  
XX CC  
XX CC  
XX CC  
XX CC  
XX CC  
XX CC  
XX CC  
XX SQ

26-MAR-2002 (first entry)  
Drosophila melanogaster polypeptide SEQ ID NO 28311.  
Drosophila; developmental biology; cell signalling; insecticide; pharmaceutical.  
Drosophila melanogaster.  
WO200171042-A2.  
27-SEP-2001.  
23-MAR-2001; 2001WO-US009231.  
23-MAR-2000; 2000US-0191637P.  
11-JUL-2000; 2000US-00614150.  
(PEKE ) PE CORP NY.  
Venter JC, Adams M, Li PWD, Myers EW;  
WPI; 2001-656860/75.  
N-PSDB; ABL11276.  
New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signaling and cell-cell interactions.  
Disclosure; SEQ ID NO 28311; 21pp + Sequence Listing; English.  
The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-ABB72072). The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

Sequence 532 AA;  
Query Match 6.3%; Score 137; DB 4; Length 532;  
Best Local Similarity 22.6%; Pred. No. 0.0046;  
Matches 62; Conservative 37; Mismatches 93; Indels 82; Gaps 10;

Qy 116 DEEDDDGDLSDIADDEDDVSYSSVDD-----VGADYEDYDTMLNKLNA-----HTG 165  
Db 153 DEEDDD-----DDDDDEEASTGRSRRRNGVAVSSGSRNVTSSSAAQRRG 204  
Qy 166 TPTSETTAEGEGETDSASSAND-----NVFDDF 196  
Db 205 KSSSEEEETSAQATSDASSDEERAQNSDSSPQOTRRARQRLSDSDSNDSE 264  
Qy 197 TSYNAHKKQKRSIADVNEEQNTQGNHTELOEKSSNEATSKESPAPLHRR----- 251  
Db 265 DSYNPNEGRRRRPGAA---RPSNRQTNGHSSKRRRLNKSANAPGSPSPRRTRIA 321  
Qy 252 -----RMHSHR-----RHLLVRKARSEDSR-----PAAHFLHSRRRHQS 287  
Db 322 GSSAEAAHSQDEHSTQDSQIGRRKKGSSRSRSGNGPSTSAANAANSLESPSRNTRA 381  
Qy 288 MGYHGMVYIGNDNERNYSQGHFQFTRDGLVTVNT 321  
Db 382 NGTTSHTLGAEND-HSY--HLPVRNGRIETSDT 412

ABB65698  
ID ABB65698 standard; protein; 2175 AA.  
XX AC  
XX ABB65698;  
XX DT  
XX DE  
XX DE  
XX KW  
XX KW  
XX OS  
XX PN  
XX PD  
XX PF  
XX PR  
XX PR  
XX PA  
XX PI  
XX DR  
XX DR  
XX PT  
XX PT  
XX PS  
XX PS  
XX CC  
XX CC  
XX CC  
XX CC  
XX CC  
XX CC  
XX CC  
XX CC  
XX CC  
XX CC  
XX SQ

26-MAR-2002 (first entry)  
Drosophila melanogaster polypeptide SEQ ID NO 23886.  
Drosophila; developmental biology; cell signalling; insecticide; pharmaceutical.  
Drosophila melanogaster.  
WO200171042-A2.  
27-SEP-2001.  
23-MAR-2001; 2001WO-US009231.  
23-MAR-2000; 2000US-0191637P.  
11-JUL-2000; 2000US-00614150.  
(PEKE ) PE CORP NY.  
Venter JC, Adams M, Li PWD, Myers EW;  
WPI; 2001-656860/75.  
N-PSDB; ABL09801.  
New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signaling and cell-cell interactions.  
Disclosure; SEQ ID NO 23886; 21pp + Sequence Listing; English.  
The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-ABB72072). The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

Sequence 2175 AA;  
Query Match 6.3%; Score 137; DB 4; Length 2175;  
Best Local Similarity 21.8%; Pred. No. 0.035;  
Matches 64; Conservative 45; Mismatches 111; Indels 74; Gaps 9;

Qy 21 PAKATSTATAQRTROLPIVLGFIGLVVAILALTIMQTRVSHLDKELSLKRVVDN 80  
Db 227 PAATVTAAGAAAAAATPATGNVKSSTTSNAHT---NSNNHQDE----- 273  
Qy 81 LQRLGINYLDEFDEKVEYNALIDYPKVGLTDEEDDDGDLSDIADDEDDVSY 140  
Db 274 -----ELDDDEEEDDE-----DEDEENASMQSNADDMELDAQOE 311  
Qy 141 SVDDVGA-----DYEDYD-----MLNKLNAHTGTTTSETTAEGEGETDS-- 182  
Db 312 TRTEPATQQOHOQQQDTELEENKAGEASLVNHNHTDSSNCSRKNNGGNESEQ 371  
Qy 183 --ASSASNDNVDPDFTSYNAHKKQKRSRSIADVNEEQNTQGNHTELOEKSSNB--- 237  
Db 372 HVASSAEDDDCANNTNTSNNNTSNTATSTNT---NNNNNNSSSGNSEKRRKKNNNNG 429  
Qy 238 -----ATSKESPAPLHRRRHHRRHLLVRKARSEDSRPAAPHLSRRRH 284  
Db 430 QPAVLLAAKDKETKALLDELQLRQAQEQTHLQIRLEE-----HLEVKRQH 476

RESULT 14  
AAU21828  
ID AAU21828 standard; protein; 330 AA.  
XX  
AC AAU21828;  
XX  
DT 06-DEC-2001 (first entry)  
XX  
DE Novel human neoplastic disease associated polypeptide #261.  
XX  
KW Human; neoplastic disease associated polypeptide; cancer;  
KW hyperproliferative disorder; neural disorder; immune system disorder;  
KW muscular disorder; reproductive disorder; gastrointestinal disorder;  
KW pulmonary disorder; cardiovascular disorder; renal disorder;  
KW neuroprotective; cytostatic; anti inflammatory; vasotropic.  
XX  
OS Homo sapiens.  
XX  
PN WO200155163-A1.  
XX  
PD 02-AUG-2001.  
XX  
PF 17-JAN-2001; 2001WO-US001358.  
XX  
PR 31-JAN-2000; 2000US-0179065P.  
PR 04-FEB-2000; 2000US-0180628P.  
PR 24-FEB-2000; 2000US-0184664P.  
PR 02-MAR-2000; 2000US-0186350P.  
PR 16-MAR-2000; 2000US-0189874P.  
PR 17-MAR-2000; 2000US-0190076P.  
PR 18-APR-2000; 2000US-0198123P.  
PR 19-MAY-2000; 2000US-0205515P.  
PR 07-JUN-2000; 2000US-0209467P.  
PR 28-JUN-2000; 2000US-0214886P.  
PR 30-JUN-2000; 2000US-0215135P.  
PR 07-JUL-2000; 2000US-0216647P.  
PR 07-JUL-2000; 2000US-0216880P.  
PR 11-JUL-2000; 2000US-0217487P.  
PR 14-JUL-2000; 2000US-0217496P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 26-JUL-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225213P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225266P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225447P.  
PR 14-AUG-2000; 2000US-0225757P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0226682P.  
PR 22-AUG-2000; 2000US-0226683P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228944P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 25-SEP-2000; 2000US-0234998P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235836P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239935P.  
PR 13-OCT-2000; 2000US-0239937P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 20-OCT-2000; 2000US-0241826P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0246474P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249246P.  
PR 17-NOV-2000; 2000US-0249265P.  
PR 17-NOV-2000; 2000US-0249297P.

```
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX
XX WPI; 2001-465558/50.
XX N-PSDB; AAS35027.
XX
XX Novel polypeptides and polynucleotides useful as diagnostic reagents to
XX diagnose diseases or disorders associated with aberrant expression or
XX activity of polypeptides, and for treating cancers, rheumatoid arthritis.
XX
XX Claim 11; SEQ ID NO 555; 687pp; English.
XX
XX The present invention relates to the isolation of novel human neoplastic
XX disease associated polypeptides, and cDNA (AAS34767-AAS35050) and DNA
XX sequences encoding for these polypeptides. The sequences of the invention
XX are useful in the diagnosis, treatment, prevention and/or prognosis of
XX disorders involving neoplastic disease such as hyperproliferative
XX disorders (e.g. leukaemia, bone cancer, bladder cancer, brain stem
XX glioma, adult liver cancer, childhood cerebellar astrocytoma, or
XX Hodgkin's lymphoma). The sequences of the invention may also be useful
XX for treating other disorders such as neural disorders, immune system
XX disorders, muscular disorders, reproductive disorders, gastrointestinal
XX disorders, pulmonary disorders, cardiovascular disorders and renal
XX disorders. The polynucleotide sequences of the invention are also useful
XX in gene therapy. AAU21568-AAU21851 represent the novel human neoplastic
XX disease associated polypeptides of the invention. Note: The sequence data
XX for this patent did not form part of the printed specification, but was
XX obtained in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX
XX SQ Sequence 330 AA;
XX
XX Query Match 6.2%; Score 135; DB 4; Length 330;
XX Best Local Similarity 21.2%; Pred. No. 0.0035;
XX Matches 68; Conservative 48; Mismatches 108; Indels 96; Gaps 12;
XX
XX Qy 20 FPAKATSTATQRTQLPIPLVIGFIGLVVA-----ILATTWQTRVSHLDKELKSL 74
XX | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
XX Db 35 FARRTSCNRCREKTEAKWKAGTETGKTLAEKSRGLFSANDWQCKTCSNVNWARRE 94
XX | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
XX Qy 75 KRVD-----NLQRLG-----INLID-----EPDEFQKEYENALIDYPKKVDG 113
XX | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
XX Db 95 CNMCTPKYAKLEERTGYGGFNRENVEYIEREESDEGEYDFGKKKK-----YRGKAVG 150
XX | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
XX Qy 114 ---LTDEDDDDGGLSDIADDDDDSVSSVDVGDVGYEDYDMLNKLNNHAHTGTTPTS 170
XX | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
XX Db 151 PASTLKEVEDKESGEE---EDEDEDLSKYKD----- 180
XX | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
XX Qy 171 ETTAAGEGETDASASNDNDVDFDTSYNAHKKQEKRSRIADVRNEEQNIQNHTEL 230
XX | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
XX Db 181 ---EDEDADLSKYN---LDASEEDSNKKSNRRSR----- 214
XX | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
XX Qy 231 QEKSSNEATSKESAPLHRRMRHSRHLHLVRKARS-----EDSRPAAHFLSSRRHQ 285
XX | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
XX Db 215 KSRSHSRSSRSSPSSSRSSRSRSSRSRSSRSRSSRSRSGRSKRSRSGRSSRSHR 274
XX | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
XX
```

```
Qy 286 GSMGYHGDYIGNDN--ERN 303
||| : : |||
Db 275 GSSSPRKRSYSSSSSSPERN 294
||| : : |||

RESULT 15
AAU86946
ID AAU86946 standard; protein; 330 AA.
XX
XX AC AAU86946;
XX
XX DT 21-MAY-2002 (first entry)
XX
XX DE Human DNA repair and processing protein, SEQ ID No 34.
XX
XX KW Human; secreted DNA repair and processing protein; immune disorder;
XX autoimmune disease; human immunodeficiency virus; HIV; infection;
XX anaemia; rheumatoid arthritis; multiple sclerosis; cancer; melanoma;
XX hyperproliferative disorder; neurological disease; Alzheimer's disease;
XX Parkinson's disease; cardiovascular disorder; cerebrovascular disorder;
XX angina; ocular disorder; corneal infection; wound healing;
XX biorythm transplantation; mesodermal tissue differentiation;
XX embryonic stem cell; catabolism; anabolism; energy storage; mental state;
XX hormone; appetite; reproductive potential; memory; stress; food additive;
XX nutrition.
XX
XX OS Homo sapiens.
XX
XX PN WO200155305-A2.
XX
XX PD 02-AUG-2001.
XX
XX PF 17-JAN-2001; 2001WO-US0001303.
XX
XX PR 31-JAN-2000; 2000US-0179065P.
XX PR 04-FEB-2000; 2000US-0180628P.
XX PR 24-FEB-2000; 2000US-0184664P.
XX PR 02-MAR-2000; 2000US-0186350P.
XX PR 16-MAR-2000; 2000US-0189874P.
XX PR 17-MAR-2000; 2000US-0190076P.
XX PR 18-APR-2000; 2000US-0198123P.
XX PR 19-MAY-2000; 2000US-0205515P.
XX PR 07-JUN-2000; 2000US-0209467P.
XX PR 28-JUN-2000; 2000US-0214886P.
XX PR 30-JUN-2000; 2000US-0215135P.
XX PR 07-JUL-2000; 2000US-0216647P.
XX PR 07-JUL-2000; 2000US-0216880P.
XX PR 11-JUL-2000; 2000US-0217487P.
XX PR 11-JUL-2000; 2000US-0217496P.
XX PR 14-JUL-2000; 2000US-0218290P.
XX PR 26-JUL-2000; 2000US-0220963P.
XX PR 26-JUL-2000; 2000US-0220964P.
XX PR 14-AUG-2000; 2000US-0224518P.
XX PR 14-AUG-2000; 2000US-0224519P.
XX PR 14-AUG-2000; 2000US-0225213P.
XX PR 14-AUG-2000; 2000US-0225214P.
XX PR 14-AUG-2000; 2000US-0225266P.
XX PR 14-AUG-2000; 2000US-0225267P.
XX PR 14-AUG-2000; 2000US-0225268P.
XX PR 14-AUG-2000; 2000US-0225268P.
XX PR 14-AUG-2000; 2000US-0225270P.
XX PR 14-AUG-2000; 2000US-0225447P.
XX PR 14-AUG-2000; 2000US-0225757P.
XX PR 14-AUG-2000; 2000US-0225758P.
XX PR 14-AUG-2000; 2000US-0225759P.
XX PR 18-AUG-2000; 2000US-0226279P.
XX PR 22-AUG-2000; 2000US-0226681P.
XX PR 22-AUG-2000; 2000US-0226868P.
XX PR 22-AUG-2000; 2000US-0227182P.
XX PR 23-AUG-2000; 2000US-0227009P.
XX PR 30-AUG-2000; 2000US-0228924P.
XX PR 01-SEP-2000; 2000US-0229287P.
XX PR 01-SEP-2000; 2000US-0229343P.
XX PR 01-SEP-2000; 2000US-0229344P.
```

```

PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231411P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 13-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 20-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 17-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.

PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 11-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.

XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
PI WPI; 2001-541497/60.
XX N-PSDB; ABK43159.
DR
XX
XX Nucleic acid molecules encoding human secreted DNA repair and processing
PT proteins, used in preventing, treating or ameliorating a disorder, e.g.
PT Alzheimer's and Parkinson's diseases and cancers.
XX
PS Claim 11; SEQ ID NO 34; 458pp; English.
XX
XX The invention relates to novel isolated nucleic acid molecules (I)
CC encoding human secreted DNA repair and processing proteins (II). (I) and
CC (II) are used to prevent, treat or ameliorate a medical condition e.g. in
CC humans. (I) and (II) may be used in the prevention, treatment and
CC diagnosis of diseases associated with inappropriate expression of (II),
CC e.g. (I), and (II) may be used to treat disorders associated with
CC decreased expression of (II) by rectifying mutations or deletions in a
CC patient's genome that affect the activity of (II). (I) and complementary
CC sequences may also be used as DNA probes in diagnostic assays (e.g.
CC polymerase chain reactions (PCR)) to detect and quantitate the presence
CC of similar sequences in samples, and so which patients may be in need of
CC restorative therapy. (II) may also be used as antigens in the production
CC of antibodies and in assays to identify modulators (agonists and
CC antagonists) of the expression and activity of (II). The anti-(II)
CC antibodies and antagonists may also be used to down regulate expression
CC and activity of (II). The anti-(II) antibodies may also be used as
CC diagnostic agents for detecting the presence of (II) in samples. The
CC disorders include e.g. immune/autoimmune diseases (e.g. human
CC immunodeficiency virus (HIV)), infections, anaemia, rheumatoid arthritis
CC and multiple sclerosis, cancers and hyperproliferative disorders (e.g.
CC melanomas), neurological diseases (e.g. Alzheimer's disease and
CC Parkinson's disease) cardio/cerebrovascular disorders (e.g. cardiac
CC arrest, angina and thrombosis), infections caused by bacteria, viruses

```

Query Match 6.2%; Score 135; DB 4; Length 330;

Best Local Similarity 21.2%; Pred No. 0.0035;

Matches 68; Conservative 48; Mismatches 108; Indels 96; Gaps 12;

QY 20 FPAKATSTATQORTRQIPLVLGGLVVA-----ILALTIMQTTTRVSHLDKELKSL 74

Db 35 FARRTSCNRCRGREKTTAKWMKAGTGKTLAKSRGLFSANDQCKTCNVNWARRE 94

QY 75 KEVVD-----NLOQRLG-----INLYD-----EFDEFQKEYENALIDYPKVDG 113

Db 95 CNMCNTPKYAKLEERTGYGGFNERNVEYIEREESDGEYDFGKKKK---YRGKAVG 150







GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 13, 2004, 10:33:39 ; Search time 18 Seconds  
(without alignments)  
1173.057 Million cell updates/sec

Title: US-09-813-329-6  
Perfect score: 2162  
Sequence: 1 MTAEFLKPFITPTSANDGF.....DRNAVLREGNRSYFGIFKV 409

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents, AA.\*  
1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pdp.\*  
2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pdp.\*  
3: /cgn2\_6/ptodata/2/iaa/6A\_COMB.pdp.\*  
4: /cgn2\_6/ptodata/2/iaa/6B\_COMB.pdp.\*  
5: /cgn2\_6/ptodata/2/iaa/PTCUS\_COMB.pdp.\*  
6: /cgn2\_6/ptodata/2/iaa/backfiles.pdp.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	132	6.1	391	US-09-342-681C-2	Sequence 2, Appli
2	129	6.0	1085	US-08-431-080-28	Sequence 28, Appli
3	129	6.0	1085	US-08-938-534-28	Sequence 28, Appli
4	129	6.0	1085	US-09-345-294-28	Sequence 28, Appli
5	122.5	5.7	706	US-08-339-152A-29	Sequence 29, Appli
6	122.5	5.7	706	US-08-007-999B-4	Sequence 4, Appli
7	122.5	5.7	706	US-08-689-276A-4	Sequence 4, Appli
8	122.5	5.7	714	US-08-990-114-3	Sequence 3, Appli
9	122.5	5.7	714	US-09-241-333-3	Sequence 3, Appli
10	121.5	5.6	296	US-07-712-476A-1	Sequence 1, Appli
11	121.5	5.6	314	US-09-134-253-1	Sequence 1, Appli
12	121.5	5.6	314	US-09-206-576-2	Sequence 2, Appli
13	118.5	5.5	282	US-07-712-476A-5	Sequence 5, Appli
14	116	5.4	249	US-09-105-343A-2	Sequence 2, Appli
15	116	5.4	300	5340934-6	Patent No. 5340934
16	115.5	5.3	289	US-09-589-287B-38	Sequence 38, Appli
17	115.5	5.3	289	US-09-588-947A-38	Sequence 38, Appli
18	115.5	5.3	289	US-09-589-286A-38	Sequence 38, Appli
19	115.5	5.3	281	US-09-216-393B-8	Sequence 8, Appli
20	115	5.3	10182	US-09-134-001C-3159	Sequence 3159, Ap
21	114.5	5.3	187	US-09-396-937-8	Sequence 8, Appli
22	114.5	5.3	414	US-07-667-276A-4	Sequence 4, Appli
23	114.5	5.3	1435	US-08-568-459A-4	Sequence 4, Appli
24	114.5	5.3	1435	US-08-487-826B-4	Sequence 4, Appli
25	114.5	5.3	1435	US-08-210-288-4	Sequence 4, Appli
26	114	5.3	878	US-09-540-236-3401	Sequence 3401, Ap
27	113	5.2	226	US-08-431-080-28	Sequence 26, Appli

28	113	5.2	226	2	US-08-938-534-26	Sequence 26, Appli
29	113	5.2	226	4	US-09-345-294-26	Sequence 26, Appli
30	112.5	5.2	553	4	US-09-533-029-90	Sequence 90, Appli
31	112.5	5.2	1162	2	US-08-728-323A-2	Sequence 2, Appli
32	112.5	5.2	1162	4	US-09-298-568-2	Sequence 2, Appli
33	112.5	5.2	1162	4	US-09-410-399-2	Sequence 2, Appli
34	112.5	5.2	1850	4	US-09-620-093A-5	Sequence 5, Appli
35	112	5.2	3696	4	US-09-134-001C-5080	Sequence 5080, Ap
36	111.5	5.2	182	4	US-09-396-937-16	Sequence 16, Appli
37	111.5	5.2	266	4	US-09-589-287B-19	Sequence 19, Appli
38	111.5	5.2	266	4	US-08-879-919-24	Sequence 24, Appli
39	111.5	5.2	266	4	US-09-588-947A-19	Sequence 19, Appli
40	111.5	5.2	266	4	US-09-589-286A-19	Sequence 19, Appli
41	111	5.1	278	6	5340934-8	Patent No. 5340934
42	111	5.1	628	3	US-08-656-034-2	Sequence 2, Appli
43	110.5	5.1	487	4	US-09-386-962C-14	Sequence 14, Appli
44	110	5.1	173	4	US-09-396-937-10	Sequence 10, Appli
45	109.5	5.1	913	4	US-08-971-089-4	Sequence 4, Appli

ALIGNMENTS

RESULT 1  
US-09-342-681C-2  
; Sequence 2, Application US/09342681C  
; Patent No. 6355782  
; GENERAL INFORMATION:  
; APPLICANT: Zonana et al.  
; TITLE OF INVENTION: Hypohyrotic ectodermal dysplasia genes and proteins  
; FILE REFERENCE: 52978  
; CURRENT APPLICATION NUMBER: US/09/342,681C  
; PRIOR FILING DATE: 1999-06-29  
; PRIOR APPLICATION NUMBER: 60/092,279  
; PRIOR FILING DATE: 1998-07-09  
; PRIOR APPLICATION NUMBER: 60/112,366  
; PRIOR FILING DATE: 1998-12-15  
; NUMBER OF SEQ ID NOS: 123  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 391  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-342-681C-2

Query Match	6.1%;	Score 132;	DB 4;	Length 391;
Best Local Similarity	28.6%;	Pred. No. 0.00046;		
Matches	44;	Conservative	31;	Mismatches 63; Indels 16; Gaps 8;
Qy	264	KARSEDSPAAHFLSSRRRHQGS-MGYHGDWYTG---NDNER---NSYQGHFQTRDGVLT	317	
Db	240	KAGTRENQPAV-VHLOG---QGSIAQVKNLDSGVLNDSRITMNPVKVPLHPRSGELE	294	
Qy	318	VTNTGLYVYVYQAQ--ICYNNSHDQNGFTVFGDTPFLQCLNTVPTNMPHKVHTHTSGLIH	375	
Db	295	VLVDGTGYFYSQVEVYINFTDFASVEVWVDEKPFQCTRSIETGKTN-YNTCYTAGVCL	353	
Qy	376	LERNERIHLDIHNDRNAVLREGNRSYFGIFKV	409	
Db	354	LKARQKIARVKNVHADIS--INMSKHTTFFGAIRL	385	

RESULT 2  
US-08-431-080-28  
; Sequence 28, Application US/08431080  
; Patent No. 5698686  
; GENERAL INFORMATION:  
; APPLICANT: Gottschling, Daniel E.  
; APPLICANT: Singer, Miriam S.  
; TITLE OF INVENTION: Telomerase Compositions and Methods  
; NUMBER OF SEQUENCES: 32  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Arnold, White & Durkee



RESULT 4  
US-09-345-294-28  
; Sequence 28, Application US/09345294  
; Patent No. 6387619  
; GENERAL INFORMATION:  
; APPLICANT: Gottschling, Daniel E.  
; SINGER, Miriam S.  
; TITLE OF INVENTION: Telomerase Compositions and Methods  
; NUMBER OF SEQUENCES: 32  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Arnold, White & Durkee  
; STREET: P.O. Box 4433  
; CITY: Houston  
; STATE: TEXAS  
; COUNTRY: UNITED STATES OF AMERICA  
; ZIP: 77210  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS/ASCII  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/345,294  
; FILING DATE: 30-Jun-1999  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/431,080  
; FILING DATE: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Parker, David L.  
; REGISTRATION NUMBER: 32,165  
; REFERENCE/DOCKET NUMBER: ARCD:155/PAR  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (512) 418-3000  
; TELEFAX: (713) 789-2679  
; TELEX: 79-0924  
; INFORMATION FOR SEQ ID NO: 28:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1085 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; SEQUENCE DESCRIPTION: SEQ ID NO: 28:  
US-09-345-294-28

Query Match 6.0%; Score 129; DB 4; Length 1085;  
Best Local Similarity 19.9%; Pred. No. 0.0039;  
Matches 73; Conservative 48; Mismatches 126; Indels 120; Gaps 12;  
Qy 69 KELSKLRVVDNLOQLRGINVLDFDEPQ-----KEYENALIDYPKVGDLTDEDDDDG 123  
Db 94 KSKLIRQINDEG-----TSSDYQAVTDGESENEEESEEDDDDDDDDD 147  
Qy 124 DGLDSIADDEDDVSYSDVDVGADYED-----YDMLNKLN 160  
Db 148 DGSDDSDSE-----TSSDDENIDFVKLTARKKRAMKALSAMNTNLTLYSSKNSNK 201  
Qy 161 NAHTGTTPTSTTAE-----GEGETDSASSANDNVPDF 196  
Db 202 NKSVKLSPKNEEEFEKEEKEEKEEQQOQESNKKYVNGSGTTTQQALSFKPKEDDG 261  
Qy 197 TSY-NAHKKKQERKSRISADVRNEQN-----IQGNHTLEQKSSNEATSKESP 244  
Db 262 ISFGNGEGYEDIGEEVDLKNENNGNEEDKLDKVMNLGNDELPPNISEDESE-- 319  
Qy 245 APLHRRMRHRRHLLVRKARSDSRPAAPHFLSSRRRHQSGMVG-----HGDWYIGND 299  
Db 320 -----YDIDQDAFVDVINNEDSHGEICTDLETGEDDLPLEE 356  
Qy 300 NERN-----SYOG--HFQTRDGLVLTWNTGLYYVYAO-ICYNNSHDQNGFIVQ 345  
Db 357 EEQNIIVSELQNDLSFDGSIHEGSDFEVDAENKFLQNEYNQENGVDEEDDEDEIMSD 416

Qy 346 GDTFFFLQ 352  
Db 417 FDMPEFYE 423  
RESULT 5  
US-08-339-152A-29  
; Sequence 29, Application US/08339152A  
; Patent No. 5643726  
; GENERAL INFORMATION:  
; APPLICANT: Tanzi, Rudolph E.  
; APPLICANT: Kovacs, Dora M.  
; TITLE OF INVENTION: Methods For Modulating Transcription  
; NUMBER OF SEQUENCES: 33  
; TITLE OF INVENTION: From The Amyloid -Protein Precursor (APP) Promoter  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
; STREET: 1100 New York Ave., NW, Suite 600  
; CITY: Washington  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20005  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/339,152A  
; FILING DATE: 10-NOV-1994  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Steffe, Eric K.  
; REGISTRATION NUMBER: 36,688  
; REFERENCE/DOCKET NUMBER: 0609.4120000  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-371-2600  
; TELEFAX: 202-371-2540  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 29:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 706 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; US-08-339-152A-29

Query Match 5.7%; Score 122.5; DB 1; Length 706;  
Best Local Similarity 21.1%; Pred. No. 0.0084;  
Matches 63; Conservative 36; Mismatches 105; Indels 95; Gaps 13;  
Qy 91 DEDEFOKEYENALIDY-----PKKVD-----GLTDEEDDDGDLDSIADDEDDV 137  
Db 218 EEDEEEEEEDEEDYDVYKSEFPTEADLEDTEAAVDEDDDEDEEEGVV---EDRDY 274  
Qy 138 SYSSVDVVGADYEDYTDMLNKLNNNAHTGTTPTSETTAEGE-----GETDSASSAND 189  
Db 275 YDTEFK--GDYNEE---NPTPEGSDGTMSDKXEITHDVKVPPTPLPTNDVDVVFETSD 328  
Qy 190 DNVFDDPTS-----YNAHKKKQER-----KSR 212  
Db 329 DNEHARFQKAKEQLIERHNRMDRVKKEWEAEALQAKNLPKAEQRTLIQHFOAMVKALE 388  
Qy 213 IADVRNEEQIQGNHTLEQKSSNE-----ATSKESPALHHR----- 250  
Db 389 KAAASEKQQLVETHLARVEAMLNRRMALENYLAALQRPDPHILQPLRYVRAEN 448  
Qy 251 -RRMHS--RH--RHLLV---RKARSEDSPAAHFLSSRRRHQSGMVGHDWYIGNDNERN 303  
Db 449 KDLRLTIRHYQHVLAVDPKAAQMSQVMTHLHVIBERRNQLSLLYKDPVVARIQEND 507  
RESULT 6  
US-08-007-999B-4



```
Db 336 DLAUVVVRTGTNRKFGYVDFESAED 360

RESULT 9
US-09-241-333-3
; Sequence 3, Application US/09241333
; Patent No. 6313266
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Yue, Henry
; APPLICANT: Corley, Neil C.
; APPLICANT: Shah, Purvi
; TITLE OF INVENTION: HUMAN NUCLEOLIN-LIKE PROTEIN
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/241,333
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/990,114
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0451 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 714 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 128842
US-09-241-333-3

Query Match 5.7%; Score 122.5; DB 4; Length 714;
Best Local Similarity 20.0%; Pred. No. 0.0086;
Matches 53; Conservative 40; Mismatches 109; Indels 63; Gaps 8;

Qy 98 KEYENALIDYPKKVDGLTDEEDDDGDLSDIADDEDDVSVSVDDVGADYEDYDMLN 157
Db 134 KNGKNA-----KKEDSDEDEDDDDSDDEDEDEED-----EFEP----- 171
Qy 158 KLNNAHTGTTPTSETTAEGEGETDSASSANDNDNVDFDFTSYNA-----HKKKQER-- 208
Db 172 PVVKGKGKVAAPASEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDE 231
Qy 209 --KRSIADVNEEQNIQGNHTELOEKSNEATSK-----SPAPLHRRRHSHRHH 259
Db 232 PVKAKNVAEDDDDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDE 291
Qy 260 LLVKAQSEDSRPAAHFLSLRRRHQSGMGYHGMVIGNDNERN-----YQGHFQTR 312
Db 292 PEAKKQKVEGSESTTFPNL-----FIGNLNPNKSVAEKVAISEPFAKN 335
Qy 313 DGLVTVTNTGLYYVYAQICYNNSHD 337
```







Db 212 VSGLLALRPGSSLRIRTL---PWAHLKAAPELTYFGLFOV 248

RESULT 15

5340934-6  
 ; Patent No. 5340934  
 ; APPLICANT: TERMINE, JOHN D.; YOUNG, MARIAN F.; FISHER, LARRY W.  
 ; ROBEY, PAMELA G.  
 ; TITLE OF INVENTION: CDNA SEQUENCES OF HUMAN BONE MATRIX PROTEINS  
 ; NUMBER OF SEQUENCES: 13  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/07/432,044  
 ; FILING DATE: 03-NOV-1989  
 ; SEQ ID NO: 6:  
 ; LENGTH: 300  
 5340934-6

Query Match 5.4%; Score 116; DB 6; Length 300;  
 Best Local Similarity 23.1%; Pred. No. 0.0099;  
 Matches 75; Conservative 32; Mismatches 101; Indels 116; Gaps 17;  
 QY 41 VLGFITGLVAVAILALTIWQTRVSHLDKELKSLKRVVDNLQORLGINYLDEF-----D 94  
 Db 5 VICFLLGLGTCAI---PVKQADSGSSEKQLYN-----KYPDAVATWLNPD 47  
 QY 95 EFQKEYENALIDYPKK-----VDGLTDEEDDDDDGGLDSI---ADDEDD-DVSYSSV 142  
 Db 48 PSQKQNLAPQTLPSKSNESHDMDDDEDDDDHVDSDQSDIDSNDSDVDVDDTDYSHQSD 107  
 QY 143 DDVGADYED--YTDMLNKL--NNAHTGTTPTSETTAEGEET----- 180  
 Db 108 ESHHSDSEDELVTFPTDLPATEVFPVPTVD-TYDGRGDSVYVGLRSKSKFRFPDIQ 166  
 QY 181 -----DSASSANDDNVFDFTSYNAHK 203  
 Db 167 YPDATDEITSHMESEELNGAYKAIPVAQDINAFSDWDSRGKDSYETSQDLD-QSAETHS 225  
 QY 204 KQQRKRSIADVRNEEQNIQNHTELQEKSSNEATSKESPAPLHRRRMRHSHRLLV- 262  
 Db 226 HKQRLYKKA--NDESN---EHSVDVIDSQELSKVSRE-----FHSHEFHS-HEDMLVV 273  
 QY 263 -RKARSDSRPAAHFHLSSRRHQ 285  
 Db 274 DPKSKEEDK-----HLKFRISHE 291

Search completed: September 13, 2004, 10:36:16  
 Job time : 20 secs

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	2162	100.0	409	9	US-09-813-329-6	Sequence 6, Appli
2	2082.5	96.3	406	9	US-09-813-329-4	Sequence 4, Appli
3	2045	94.6	409	9	US-09-813-329-2	Sequence 54, Appli
4	149	6.9	27	9	US-09-813-329-54	Sequence 54, Appli
5	149	6.9	27	9	US-09-813-329-64	Sequence 64, Appli
6	145.5	6.7	511	16	US-10-437-963-148275	Sequence 148275
7	141	6.5	428	16	US-10-437-963-148275	Sequence 148275
8	140.5	6.5	406	16	US-10-437-963-154595	Sequence 154595
9	137	6.5	699	16	US-10-408-765A-434	Sequence 434, Appli
10	135	6.2	330	9	US-09-764-848-34	Sequence 34, Appli
11	135	6.2	330	14	US-10-116-016-34	Sequence 34, Appli
12	135	6.2	330	14	US-10-103-313-555	Sequence 555, Appli
13	135	6.2	330	14	US-10-232-020-34	Sequence 34, Appli
14	132	6.1	391	9	US-09-813-329-8	Sequence 8, Appli
15	132	6.1	391	10	US-09-739-658B-2	Sequence 2, Appli



Qy	108	PKKVDGLTDEEDDDGGDGLSDIADDEDDVSYSVVDVGADYEDYDTMLNKLNNAHGTGTT	167
Db	173	PKKV-ASSDKGEEDSKGSESDESDSLSDSKSD-----DTRRK-KKGRKGSH	221
Qy	168	PTSET-----TAEGEGETDSASSANDDNVDFDFTSYNAHKKKOERKSRSTADVRNE	219
Db	222	RSKRSRHRHRHSHSDTEGD-DNSKAEDSEGSVDSDESDMDRRKKKRSR-----RHK	272
Qy	220	EQNIQNHTELQEKSSNE-----ATSKESPAPLHHRHRHMHSHRHLVRKAR	266
Db	273	KSKRRGRSSRRKKRKNDSNTASEGSGSSEEEAVALAASGSPSPLRDSKKKSSRRKRSKQSD	332
Qy	267	SEDSRPA 273	
Db	333	SEDQAPS 339	
RESULT 7			
US-10-451-467A-448			
; Sequence 448, Application US/10451467A			
; Publication No. US20040161840A1			
; GENERAL INFORMATION:			
; APPLICANT: CONTRERAS, ROLAND HENRI			
; APPLICANT: EBERHARDT, INES			
; APPLICANT: LUYTEN, WALTER HERMAN MARIA LOUIS			
; APPLICANT: REEKWANS, RIEKA JOSEPHINA			
; TITLE OF INVENTION: BAX-RESPONSIVE GENES FOR DRUG TARGET IDENTIFICATION IN			
; TITLE OF INVENTION: YEAST AND FUNGI			
; FILE REFERENCE: JAB-1667			
; CURRENT APPLICATION NUMBER: US/10/451,467A			
; CURRENT FILING DATE: 2003-06-19			
; PRIOR APPLICATION NUMBER: EP 00870318.3			
; PRIOR FILING DATE: 2000-12-22			
; PRIOR APPLICATION NUMBER: EP 01870002.1			
; PRIOR FILING DATE: 2001-01-04			
; PRIOR APPLICATION NUMBER: EP 01870003.9			
; PRIOR FILING DATE: 2001-01-09			
; NUMBER OF SEQ ID NOS: 732			
; SOFTWARE: Patentin version 3.1			
; SEQ ID NO 448			
; TYPE: PRT			
; ORGANISM: Candida albicans			
US-10-451-467A-448			
Query Match 6.5%; Score 141; DB 16; Length 428;			
Best Local Similarity 20.8%; Pred. No. 0.0063;			
Matches 76; Conservative 68; Mismatches 138; Indels 84; Gaps 16;			
Qy	36	QLIPLV-LGFIGLGVVAILALTITWQTRVSHLDKEL-----KSLKRVVDNL	81
Db	66	QLLPITAKMSSNTQDLVLAYINDVYSRNEELSKLALKSLACKELPKVSKQLESIIDEV	125
Qy	82	QOR-----LGINYLDFEFQKEYENALIDVPKK---VDGLTDEEDDDGGDGLSDIADDED	134
Db	126	ENQEKSKPRNSSDSDSSSESSTSDSSSSSDSDSSSDSSSDSSSDSSSDSSSDSSD	185
Qy	135	DDVSYSSVD-DVGADYEDYDTMLN-KLNNAHGTGTTPTSETTAGEGETDSASSANDDNV	192
Db	186	SDDEEDKEDKEAKDNKDSEDSENEKVEEDNKDTSDSSSSSKSDSDSDSSSDSSSS	245
Qy	193	FDDFTSYNAHKKKOERKSRSTADVRNE---EQNIQNHTEL---QEKSSNEATSKESPAP	246
Db	246	DSDDSSSDSSSDSSSDSSSDSDSDSDSDSDSDSDSDSDSDSDSDSSSDSSSDSS	305
Qy	247	LHHRHRHMSR--HRHLLVRKARSE-----DSRPAA-----HFLSSRRR--	283
Db	306	PEDKKRKHDTDDIKEEKPKVKFKNESESSASSSTSDSIPTAPELPKQQRKHFSRIDRSKV	365
Qy	284	-----HQSGMGVHGD-----YTGNDN--BRNSYQGHFQTRDGLVLT	318
Db	366	NPENSVLODNTYKGAAGTWGEKAKLLQVRGKDFTKNKNMKRGYSYGK-----GSITL	419







;  
; PRIOR FILING DATE: 2000-07-07  
; PRIOR APPLICATION NUMBER: 60/225,267  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/216,880  
; PRIOR FILING DATE: 2000-07-07  
; PRIOR APPLICATION NUMBER: 60/225,270  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/251,869  
; PRIOR FILING DATE: 2000-12-08  
; PRIOR APPLICATION NUMBER: 60/235,834  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: 60/234,274  
; PRIOR FILING DATE: 2000-09-21  
; PRIOR APPLICATION NUMBER: 60/234,223  
; PRIOR FILING DATE: 2000-09-21  
; PRIOR APPLICATION NUMBER: 60/228,924  
; PRIOR FILING DATE: 2000-08-30  
; PRIOR APPLICATION NUMBER: 60/224,518  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/236,369  
; PRIOR FILING DATE: 2000-09-29  
; PRIOR APPLICATION NUMBER: 60/224,519  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/220,964  
; PRIOR FILING DATE: 2000-07-26  
; PRIOR APPLICATION NUMBER: 60/241,809  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/249,299  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/236,327  
; PRIOR FILING DATE: 2000-09-29  
; PRIOR APPLICATION NUMBER: 60/241,785  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/244,617  
; PRIOR FILING DATE: 2000-11-01  
; PRIOR APPLICATION NUMBER: 60/225,268  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/236,368  
; PRIOR FILING DATE: 2000-09-29  
; PRIOR APPLICATION NUMBER: 60/251,856  
; PRIOR FILING DATE: 2000-12-08  
; PRIOR APPLICATION NUMBER: 60/251,868  
; PRIOR FILING DATE: 2000-12-08  
; PRIOR APPLICATION NUMBER: 60/229,344  
; PRIOR FILING DATE: 2000-09-01  
; PRIOR APPLICATION NUMBER: 60/234,997  
; PRIOR FILING DATE: 2000-09-25  
; PRIOR APPLICATION NUMBER: 60/229,343  
; PRIOR FILING DATE: 2000-09-01  
; PRIOR APPLICATION NUMBER: 60/229,345  
; PRIOR FILING DATE: 2000-09-01  
; PRIOR APPLICATION NUMBER: 60/229,287  
; PRIOR FILING DATE: 2000-09-01  
; PRIOR APPLICATION NUMBER: 60/229,513  
; PRIOR FILING DATE: 2000-09-05  
; PRIOR APPLICATION NUMBER: 60/231,413  
; PRIOR FILING DATE: 2000-09-08  
; PRIOR APPLICATION NUMBER: 60/229,509  
; PRIOR FILING DATE: 2000-09-05  
; PRIOR APPLICATION NUMBER: 60/236,367  
; PRIOR FILING DATE: 2000-09-29  
; PRIOR APPLICATION NUMBER: 60/237,039  
; PRIOR FILING DATE: 2000-10-02  
; PRIOR APPLICATION NUMBER: 60/237,038  
; PRIOR FILING DATE: 2000-10-02  
; PRIOR APPLICATION NUMBER: 60/236,370  
; PRIOR FILING DATE: 2000-09-29  
; PRIOR APPLICATION NUMBER: 60/236,802  
; PRIOR FILING DATE: 2000-10-02  
; PRIOR APPLICATION NUMBER: 60/237,037  
; PRIOR FILING DATE: 2000-10-02  
; PRIOR APPLICATION NUMBER: 60/237,040  
; PRIOR FILING DATE: 2000-10-02

;  
; PRIOR APPLICATION NUMBER: 60/240,960  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/239,935  
; PRIOR FILING DATE: 2000-10-13  
; PRIOR APPLICATION NUMBER: 60/239,937  
; PRIOR FILING DATE: 2000-10-13  
; PRIOR APPLICATION NUMBER: 60/241,787  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/246,474  
; PRIOR FILING DATE: 2000-11-08  
; PRIOR APPLICATION NUMBER: 60/246,532  
; PRIOR FILING DATE: 2000-11-08  
; PRIOR APPLICATION NUMBER: 60/249,216  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,210  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/226,681  
; PRIOR FILING DATE: 2000-08-22  
; PRIOR APPLICATION NUMBER: 60/225,759  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/225,213  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/227,182  
; PRIOR FILING DATE: 2000-08-22  
; PRIOR APPLICATION NUMBER: 60/225,214  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/235,836  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: 60/230,438  
; PRIOR FILING DATE: 2000-09-06  
; PRIOR APPLICATION NUMBER: 60/215,135  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: 60/225,266  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/249,218  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,208  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,213  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,212  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,207  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,245  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,244  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,217  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,211  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,215  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,264  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,214  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,297  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/232,400  
; PRIOR FILING DATE: 2000-09-14  
; PRIOR APPLICATION NUMBER: 60/231,242  
; PRIOR FILING DATE: 2000-09-08  
; PRIOR APPLICATION NUMBER: 60/232,081  
; PRIOR FILING DATE: 2000-09-08  
; PRIOR APPLICATION NUMBER: 60/232,080  
; PRIOR FILING DATE: 2000-09-08  
; PRIOR APPLICATION NUMBER: 60/231,414  
; PRIOR FILING DATE: 2000-09-08  
; PRIOR APPLICATION NUMBER: 60/231,244  
; PRIOR FILING DATE: 2000-09-08  
; PRIOR APPLICATION NUMBER: 60/233,064

```
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/233,063
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/232,397
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/232,399
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/232,401
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/241,808
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,826
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,786
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,221
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/246,475
; PRIOR FILING DATE: 2000-11-08

Query Match          6.2%; Score 135; DB 14; Length 330;
Best Local Similarity 21.2%; Pred. No. 0.014;
Matches 68; Conservative 48; Mismatches 108; Indels 96; Gaps 12;

Qy 20 FPAKATSTATQRTQRIPLVLGFIGLGLVA-----ILALTWQTTTVSHLDKELKSL 74
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 35 FARTSCNRCGREKTEAKMKKAGTGTEIGTLAEKSRGLFSANDWQCKTCSNVNWARRE 94
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 75 KRVD-----NLQRLG-----INLYD-----EPDEFQKEYENALIDYPKKVDG 113
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 95 CNMCNTPKYAKLEERTGYGGGFNERENVEYIEREESDGEYDFGKKKK-----YRKA 150
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 114 ---LTDEDDDDGGLDSDIADDDVSVSVDDVDGADYEDYDMLNKLNNHAHTGTTPTS 170
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 151 PASILKEVEDKESEGE---EDDEDLSKYXD-----180
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 171 ETTAAGEGETDSASASNDNVDFDFTSYNAHKKQERKSRSIADVRNEEQIQQNHTEL 230
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 181 ----EDEDADLSKN-----LDASEREEDSKKSNRRSRS-----214
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 231 QEKSSNEATKESAPLHRRMRHSRHLVLRKARS-----EDSRPAAPHLSRRRHQ 285
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 215 KRSRSHRSRSSRSPSSRSRSRSRSSSSSSQSRSSRSRGRSRSSRSRSHR 274
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 286 GSMGYGDMYIGNDN--ERN 303
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 275 GSSSPRKSSYSSSSSSPERN 294
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 14
US-09-813-329-8
; Sequence 8, Application US/09813329
; Patent No. US20020012968A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Subb Company
; TITLE OF INVENTION: No. US20020012968A1el Drosophila Tumor Necrosis Factor Class Mole
; FILE REFERENCE: Variants Thereof
; FILE REFERENCE: D0016.np
; CURRENT APPLICATION NUMBER: US/09/813,329
; CURRENT FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: 60/190,816
; PRIOR FILING DATE: 2000-03-21
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 8
; LENGTH: 391
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-813-329-8
```

Query Match 6.1%; Score 132; DB 9; Length 391;  
Best Local Similarity 28.6%; Pred. No. 0.03;  
Matches 44; Conservative 31; Mismatches 63; Indels 16; Gaps 8;

```
Qy 264 KARSEDSRPAAPHLSRRRHQGS-MGYHGDYVIG--NDNER---NSYQGHFQTRDGVLT 317
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 240 KAGTRENQPAV-VHLQ-----QGSAIQVKNDLSGVLNDMSRTMNPVKFKLHPRSGELE 294
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 318 VTNTGLYYVYQA--ICYNNSHDONGFIVFOGDTTFFLOCLNTVPTNMPHKVHTCHTSLGIH 375
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 295 VLVDGTIFYSQVEVYINFTDFASVEVVVDEKPFLOQCTRSIETGKTN-YNTCYTAGVCL 353
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 376 LERNERIHLDIHDRNAVLREGNNRSYFGIFKV 409
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 354 LKARQKIAVQVHADIS--INMSKHITFFFGAIRL 385
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

## RESULT 15

```
US-09-729-658B-2
; Sequence 2, Application US/09729658B
; Publication No. US20030023991A1
; GENERAL INFORMATION:
; APPLICANT: Zonana et al.
; TITLE OF INVENTION: Hypohydrotic ectodermal dysplasia genes and proteins
; FILE REFERENCE: 55924
; CURRENT APPLICATION NUMBER: US/09/729,658B
; CURRENT FILING DATE: 2000-12-04
; PRIOR APPLICATION NUMBER: 09/342,681
; PRIOR FILING DATE: 1999-06-29
; PRIOR APPLICATION NUMBER: 60/092,279
; PRIOR FILING DATE: 1998-07-09
; PRIOR APPLICATION NUMBER: 60/112,366
; PRIOR FILING DATE: 1998-12-15
; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 391
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-729-658B-2
```

Query Match 6.1%; Score 132; DB 10; Length 391;  
Best Local Similarity 28.6%; Pred. No. 0.03;  
Matches 44; Conservative 31; Mismatches 63; Indels 16; Gaps 8;

```
Qy 264 KARSEDSRPAAPHLSRRRHQGS-MGYHGDYVIG--NDNER---NSYQGHFQTRDGVLT 317
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 240 KAGTRENQPAV-VHLQ-----QGSAIQVKNDLSGVLNDMSRTMNPVKFKLHPRSGELE 294
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 318 VTNTGLYYVYQA--ICYNNSHDONGFIVFOGDTTFFLOCLNTVPTNMPHKVHTCHTSLGIH 375
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 295 VLVDGTIFYSQVEVYINFTDFASVEVVVDEKPFLOQCTRSIETGKTN-YNTCYTAGVCL 353
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 376 LERNERIHLDIHDRNAVLREGNNRSYFGIFKV 409
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 354 LKARQKIAVQVHADIS--INMSKHITFFFGAIRL 385
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

Search completed: September 13, 2004, 10:46:53  
Job time : 133 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 13, 2004, 10:32:13 ; Search time 19 Seconds  
(without alignments)  
2070.649 Million cell updates/sec

Title: US-09-813-329-6  
Perfect score: 2162  
Sequence: 1 MTATLKPFTTPTSDNGF.....DRNAVLRGNRSYFGIFKV 409

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues  
Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_78.\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	165.5	7.7	3268	2 S69625	hypothetical prote
2	153.5	7.1	852	2 A34373	histidine-rich cal
3	144.5	6.7	764	2 S64951	hypothetical prote
4	138.5	6.4	1192	2 A71623	probable secreted
5	138.5	6.4	1883	2 T13944	chromodomain helic
6	138.5	6.4	3097	2 T28635	glutamate synthase
7	137	6.3	699	2 A54660	histidine rich cal
8	137	6.3	2175	1 S03170	homeotic protein c
9	135	6.2	662	2 T18233	probable transcrip
10	134	6.2	466	2 S19365	hypothetical prote
11	133	6.2	2500	2 G71609	hypothetical prote
12	130.5	6.0	775	2 T21259	hypothetical prote
13	129	6.0	692	2 T03377	homeotic protein H
14	129	6.0	1085	2 S55352	IFH1 protein - yea
15	127	5.9	872	2 T18861	probable peptide-a
16	126.5	5.9	1274	2 A89959	hypothetical prote
17	126	5.8	463	2 T41390	zinc finger protei
18	125	5.8	432	2 G71621	MAK16 homolog PFB0
19	125	5.8	709	2 A96349	F8K7.2 protein - A
20	125	5.8	1672	2 T46237	hypothetical prote
21	125	5.8	2910	2 T28156	DNA-directed RNA p
22	124.5	5.8	3844	2 T18402	asparagine/asparta
23	123.5	5.7	487	2 T10215	hypothetical prote
24	123.5	5.7	630	2 H96770	protein heat shock
25	123.5	5.7	1390	2 T31353	polyprotein - Arab
26	122.5	5.7	406	2 H84590	hypothetical prote
27	122.5	5.7	713	2 A27441	nucleolin - Chines
28	122.5	5.7	1001	2 T16419	hypothetical prote
29	122.5	5.7	1379	2 S64603	YTA7 protein - yea

30	121.5	5.6	314	1 S09575	osteopontin precu
31	121.5	5.6	699	2 G86311	hypothetical prote
32	121	5.6	292	2 T30321	hypothetical prote
33	120.5	5.6	675	2 T03744	myoD protein inhib
34	120.5	5.6	964	2 S06028	gene suppressor-of
35	120.5	5.6	5138	2 B96695	hypothetical prote
36	120	5.6	294	1 A37818	osteopontin precu
37	120	5.6	6713	2 B89921	osteopontin precu
38	119.5	5.5	303	1 GEPGO	osteopontin precu
39	119.5	5.5	1078	2 T18352	protein p120 - Myc
40	119	5.5	665	2 B71609	hypothetical prote
41	119	5.5	770	2 H84463	hypothetical prote
42	118.5	5.5	1046	2 T29776	hypothetical prote
43	118.5	5.5	1105	2 T18295	Ap-3 adaptor compl
44	118	5.5	1067	2 S35423	protein kinase egg
45	117.5	5.4	290	2 G86155	hypothetical prote

ALIGNMENTS

RESULT 1

S69625  
hypothetical protein YDR457w - yeast (Saccharomyces cerevisiae)  
C:Species: Saccharomyces cerevisiae  
C:Date: 22-Aug-1996 #sequence\_revision 06-Sep-1996 #text\_change 23-Mar-2001  
C:Accession: S69625  
R:Dietrich, F.S.  
submitted to the EMBL Data Library, August 1995  
A:Description: The sequence of S. cerevisiae cosmid 9410, 8035, 8166, and 9787.  
A:Reference number: S69554  
A:Accession: S69625  
A:Molecule type: DNA  
A:Residues: 1-3268 <DIE>  
A:Cross-references: EMBL:U33050; NID:g927726; PIDN:AAB64910.1; PID:g927738; MIPS:YDR457w  
C:Genetics:  
A:Gene: SGD:TOM1  
A:Cross-references: SGD:S0002865; MIPS:YDR457w  
A:Map position: 4R

Query Match 7.7%; Score 165.5; DB 2; Length 3268;  
Best Local Similarity 19.4%; Pred. No. 0.028;  
Matches 78; Conservative 66; Mismatches 142; Indels 117; Gaps 16;

Qy	51	VAILALTIWTVRVSHLDKELSKRVVDNLQORLGI	-----NYLDEFDEFOKE-----YE 101
Db	1856	VAVEALNTISSTR	-----NNFSEHKIEDHDEVEDESDKEIPDMFK 1900
Qy	102	NALIDYPKKVDGLTD	----BEDDDD-----GLDSIADDEDDVSYSS 141
Db	1901	NSAL	-----GMVDVEDIEDDDDDTSLIGDDDAFVSDNGFEVVFSDDDDMGEED 1953
Qy	142	VDDVGADYDYDMLKNNAHTGTTTSTTAGEGETDSASASANDNVDFD	----TS 198
Db	1954	ADDAARSDEE-NELSEMSQSTADGTDVDEVDADGLIINIDQPSGDDEEMADYDANIS	2012
Qy	199	YNAHKKKQERKSIAVDRNEEQINQNHTELOE	-----KSSNEATSKES 243
Db	2013	HSSHSENEDDASMDVIEVYDDELS-SCYDVLDSDYDVDESGLSLISIDSDSESSE	2071
Qy	244	PAPLHRRRRMHSRRHLLVRKAR-SEDSRPAAPHLLSSRRRRHQSGMGYHGDMYIGNDNER	302
Db	2072	DEPINSTRMGDSRRRLIAEGVELTDDSQ	-----GESEED 2106
Qy	303	NSVQGHQTQDGLVTNTGLYYVYAIQICYNNSHDQNGFIVFQG--DTPFLQCLNTVPTN	360
Db	2107	D--RGVFRGIEHIFSNENEPLFRVHDEMRRHNRHSINRTHFSAMSAPSLSLLNRGRN	2164
Qy	361	MPHKVHTCHTSGLIHLF	-----RNERIHLKDI 387
Db	2165	QSNLINPLGTGLQEQVENDISDQTVAGSGSRPRASHLHFSEV	2207

RESULT 2	
A34373	histidine-rich calcium-binding protein precursor - rabbit
C:Species:	Oryctolagus cuniculus (domestic rabbit)
C:Date:	08-Jun-1990 #sequence_revision 08-Jun-1990 #text_change 05-Nov-1999
C:Accession:	A34373
R:Hofmann, S.L.;	Goldstein, J.L.; Orth, K.; Moomaw, C.R.; Slaughteer, C.A.; Brown, M.S.
J. Biol. Chem.	264, 18083-18090, 1989
A:Title:	Molecular cloning of a histidine-rich Ca(2+)-binding protein of sarcoplasmic re
A:Reference number:	A34373; PMID:90036884; PMID:2808365
A:Accession:	A34373
A:Status:	preliminary
A:Molecule type:	mRNA
A:Residues:	1-852 <HOF>
A:Cross-references:	GB:J05080; NID:g165099; PIDN:AAA31279.1; PID:g165100
C:Keywords:	calcium binding
Query Match	
Best Local Similarity 7.1%; Score 153.5; DB 2; Length 852;	
Matches 53; Conservative 31; Mismatches 88; Indels 61; Gaps 8;	
Qy	116 DEEDDDGGLDS-----IADEDDVSYSVDVVGADYEDYDTMLNKLNAH 163
Db	293 DEDDDDGSDTSRQAHRRHGRHEEEDDD-----DDEG-----DSTESDRHQARRH 342
Qy	164 TGTTPSTTTAEGEGETDSA-----SSASNDNVFDDFTSYNAHKKKQERKSR 211
Db	343 RGHREEDDDDDGSDTSRQAHRRHGRHEEEDDDDDGSDTSRQAHRRHGRH 402
Qy	212 SIADVNRNEQ-----NIQGNHTELQEKSSNEATSKESPAPLHRRRRMHSRHH 259
Db	403 EEEEDDDDGSTESDHQAHRRHGRHEEEDDEDDDGSDTSRQAHRRHGRHGHGEED 462
Qy	260 LLVRKARSEDSPAAHFLSRHHQSGMGVH-----GDMYIGNDNERNSTQGH 308
Db	463 -----EDDDDEGEHHVP-----HRGHRGHEEDDGGDDGDDSTENGQAH 504
RESULT 3	
S64951	hypochemical protein YLR114c - yeast (Saccharomyces cerevisiae)
N:Alternate names:	hypothetical protein L2941
C:Species:	Saccharomyces cerevisiae
C:Date:	01-Aug-1995 #sequence_revision 24-May-1996 #text_change 19-Apr-2002
C:Accession:	S64951; S69401
R:Verhaaselt, P.;	Voet, M.; Volckaert, G.
submitted to the Protein Sequence Database,	May 1996
A:Reference number:	S64943
A:Accession:	S64951
A:Molecule type:	DNA
A:Residues:	1-764 <VER>
A:Cross-references:	EMBL:Z73286; NID:g1360509; PIDN:CAA97681.1; PID:e245807; PID:g136051
A:Experimental source:	strain S286C
R:Verhaaselt, P.;	Volckaert, G.
submitted to the EMBL Data Library,	September 1995
A:Reference number:	S69393
A:Accession:	S69401
A:Molecule type:	DNA
A:Residues:	1-764 <VER>
A:Cross-references:	EMBL:X89514; NID:g1297019; PIDN:CAA61692.1; PID:e198747; PID:g129702
C:Genetics:	
A:Gene:	SGD:EFR4
A:Cross-references:	SGD:S0004104
A:Map position:	12R
Query Match	
Best Local Similarity 6.7%; Score 144.5; DB 2; Length 764;	
Matches 51; Conservative 36; Mismatches 79; Indels 41; Gaps 7;	
Qy	55 ALTIWQTTR-----YSHLDKELKSLKRVVDNLQOR---LGINYLD--EFDEF 96
Db	535 AANLWNSWKYFNKPKNTANEDVTESTEDUKNRKTSNAIQKAMMGLGLHYPDAETDQO 594

Qy	97 QKEYENALIDYPKKVDGLTDEEDDDGGLDSTADDEDDVSYSVDDVVGADYEDYTDML 156
Db	595 SEEVGNSDEDNDDTDEDDDDGCGDDDDSDDDDDDD---GEGDENGDDGEG----- 646
Qy	157 NKLNAHTGTTPTSETTAEGEGETDSASSASND-DNVFDDFTSYNA-HKKQKRSRSIA 214
Db	647 -----DENGDDGEGDENGCKEDSQDFNSGFTDVTNINTDHDKENBQNFEGNA 693
Qy	215 DVNRNEQIQGNHTELQEKSSNEATSK 241
Db	694 ENFNEDETADVADKIEGGPESNKNDSK 720
RESULT 4	
A71623	probable secreted protein PFB0115w - malaria parasite (Plasmodium falciparum)
C:Species:	Plasmodium falciparum
C:Date:	13-Nov-1998 #sequence_revision 13-Nov-1998 #text_change 21-Jul-2000
C:Accession:	A71623
R:Gardner, M.J.;	Tettelin, H.; Carucci, D.J.; Cummings, L.M.; Aravind, L.; Koonin, E.V.;
.;	Pertea, M.; Salzberg, S.; Zhou, L.; Sutton, G.G.; Clayton, R.; White, O.; Smith, H.O.;
Science	282, 1126-1132, 1998
A:Title:	Chromosome 2 sequence of the human malaria parasite Plasmodium falciparum.
A:Reference number:	A71600; PMID:99021743; PMID:9804551
A:Accession:	A71623
A:Status:	preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type:	DNA
A:Residues:	1-1192 <GAR>
A:Cross-references:	GB:AE001373; GB:AE001362; NID:g3845097; PIDN:AAC71813.1; PID:g3845098
A:Experimental source:	clone 3D7
C:Genetics:	
A:Gene:	PFB0115w
Query Match	
Best Local Similarity 6.4%; Score 138.5; DB 2; Length 1192;	
Matches 73; Conservative 60; Mismatches 145; Indels 79; Gaps 17;	
Qy	68 DKELSKLRVVDNLQORLGINYLDFEFQKEYENAL-----IDYPKKVDGLTDEE----- 118
Db	567 EKETEVEKKKTEKDEE--GTDYEEDTDDSDKDVETEVEETDAEDKEENEEGTDDSEDKVE 624
Qy	119 ----DDDDGGLDSTADDEDDVSYSVDDVVGADYEDYTDMLNNAHTGTTPTSETTA 174
Db	625 ETDLDQDEEDGEDKEDDKEDDKEDDKEDDKED-----DKEKYK 667
Qy	175 RGEGETDSASSASND--DNVFDFTSYNAHKKKQKRSIADVNRNEQIQGNHTELQE 232
Db	668 EDDKEDDKEDDKEDDKEDKED--KEDNKEKDKEDDKEDKED--DKEKDKEDN--KEKD 722
Qy	233 KSSNEATSKESPAPLHRRRRMHSRHHLLVKARSEDSRPAAHFLHSRRRRHQSGMGVHG 292
Db	723 KEDNKEKDEDD-----KEKHDKHVRRIKKRKKDDDDYDES----LKTKNYYPHNMFTGQ 772
Qy	293 DMYIG--NDNERNSTQGHFQTRDGLVTNTGLYVYVYAAQICVNNSHDQNGFTVFO----- 345
Db	773 QQYFPYNPLEQONQLHH-----IQQQNVHPHIIQQQNNHHPHIILOEKEKHP 825
Qy	346 GUTPFLQCLNTVP-----TNMPHKVHTCHT-----SGLIHLERNERIHLKDINH 390
Db	826 QGIPKEQFPYNNVPYILKKGLEPKTHN-HVKEDQPNIKQGVV---KGQEPHVDMMHNN 878
RESULT 5	
T13944	chromodomain helicase - fruit fly (Drosophila melanogaster)
C:Species:	Drosophila melanogaster
C:Date:	20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Dec-2002
C:Accession:	T13944
R:Stokes, D.G.;	Tartof, K.D.; Perry, R.P.
Proc. Natl. Acad. Sci. U.S.A.	93, 7137-7142, 1996
A:Title:	CHD1 is concentrated in interbands and puffed regions of Drosophila polytene ch
A:Reference number:	Z17823; PMID:96293489; PMID:8692958
A:Accession:	T13944

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-1883 <STO>

A:Cross-references: EMBL:L77907; NID:g1448982; PID:g1448983; PIDN:AAC37264.1

C:Genetics:

A:Gene: CHD-1

A:Cross-references: FlyBase:FBgn0016132

C:Superfamily: chromodomain helicase CHD1; chromobox homology

C:Keywords: DNA binding

Query Match 6.4%; Score 138.5; DB 2; Length 1883;  
Best Local Similarity 24.5%; Pred. No. 0.76; Mismatches 77; Indels 45; Gaps 8;  
Matches 52; Conservative 38;

Qy 110 KVDGLTDEDDDDGGLDSIADDEDDVSYSSVDVGADYDYMNLNK-LNNAHTGT- 167

Db 83 KTFGTTDQDESSDG-----SSGSDSDSDAEGPSDQRNQSNINANTSSSL 128

Qy 168 PTSETTAEGEGETDS-----ASSASND-----NVFDDFTS-----YNAHKKKQE 207

Db 129 FKPEQNEEDNETAQGOQPASDASADESSANVSPTSSSSSEEEEDYRPKRTQA 188

Qy 208 RKSRSIAD-----VRNEQNIQGNHTELOEKSNEATS--KSPAPLHRRHRRHSHR 258

Db 189 RKPPTAAEKSKAPAPKNKKTWDSDESDSDDEVSTAQKRKPATTSRSKLAQQQ 248

Qy 259 HLLVTKARSEDSRPAAHFHLSSRRRHQSGMGV 290

Db 249 RRRVKPFSSDDDDASKRCATRKGAHSVY 280

RESULT 6

T28635

glutamate synthase (NADH2) (EC 1.4.1.14) - malaria parasite (Plasmodium falciparum)

C:Species: Plasmodium falciparum

C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 03-Jun-2002

C:Accession: T28635

R:Cowan, G.M.

submitted to the EMBL Data Library, April 1998

A:Reference number: Z20490

A:Accession: T28635

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-3097 <COW>

A:Cross-references: EMBL:Y17045; NID:e1286063; PID:e1286064; PIDN:CAA76602.1

C:Genetics:

A:Gene: Glts

C:Keywords: oxidoreductase

Query Match 6.4%; Score 138.5; DB 2; Length 3097;  
Best Local Similarity 23.7%; Pred. No. 1.4;  
Matches 49; Conservative 33; Mismatches 70; Indels 55; Gaps 10;

Qy 78 VDNLQQRIGINYLDEFBQKEYENALIDYPPKVDGLTDEDDDDGGLDSIADDEDDV 137

Db 1622 IDN-DENEDMQIDIDDEDEYNN-----DGSDDSDSDSDGND-D-SDNDDDI 1670

Qy 138 -----SYSSVDVGADYEDYDM-----LNKLNNATGTPTSETTAEGEGETDS 182

Db 1671 DDLDDIDDLDDVDVD-DLDDVDLDDLDLDDVDVDDVHN-----ADNVNCRKDNHTD- 1724

Qy 183 ASSASNDNVFDDFTSYN-----AHKKK-----QERKSRSIADVNEE 220

Db 1725 -HDYDDNNVXDEFPNNDKSGQNSREKRNKKSKNRNIIYKYQKEKGSYNINKKHN 1783

Qy 221 QNIQGNHTELOEKSNEATSKEAPL 247

Db 1784 NNINKNTKTKKKIHLKKVKPIPI 1810

RESULT 7

A54660

histidine rich calcium binding protein - human

C:Species: Homo sapiens (man)

C:Date: 02-Jun-1995 #sequence\_revision 02-Jun-1995 #text\_change 05-Nov-1999

C:Accession: A54660

R:Hofmann, S.L.; Topham, M.; Heieh, C.L.; Francke, U.

Genomics 9, 656-669, 1991

A:Title: cDNA and genomic cloning of HRC, a human sarcolemmal reticulum protein, and 10

A:Reference number: A54660; MUID:91244309; PMID:2037293

A:Accession: A54660

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-699 <HOF>

A:Cross-references: GB:M60052; NID:g183918; PIDN:AAA88071.1; PID:g183919

C:Genetics:

A:Gene: GDB:HRC

A:Cross-references: GDB:126369; OMIM:142705

A:Map position: 19q13.3-19q13.3

C:Keywords: calcium binding

Query Match 6.3%; Score 137; DB 2; Length 699;  
Best Local Similarity 20.7%; Pred. No. 0.29;  
Matches 60; Conservative 36; Mismatches 116; Indels 76; Gaps 10;

Qy 91 DEFDEFQKEYENALIDYPPKVD---GLTDEDDDDG------GLDSTADDEDD 135

Db 193 EEEEEEEEEEAETEGHQARRHRCGSEDEVDSDGHHHGHSHRHQGHHEEDDDDD 252

Qy 136 DVSYSSVDDVGADYEDYDMLNKLNNATG-TTPTSETTAEGEGETDSA---SSASNDN 191

Db 253 DDDDDDDDDVSYEYRQA-----HRHQHGIEEDVDSDGHHHHRDPHRRHSHEEDN 305

Qy 192 VFDDFTSYNAHK--KKQERKSRSIADVRNE-----PQNIQGNHTELOEKSNEATS 243

Db 306 DDDVSTEGYQAHRRHQDRKEVEAVSGEHHHVPDRHQGHGHRDEEDEDVSTERWHQ 365

Qy 244 PAPLHH-----RRMHSRHRHLVVRK 264

Db 366 PQVHHGLVDDEEEEEEITVQGHYVASHQPRGHKSDEEDFQDEYKTEVPHHHHRVPRE 425

Qy 265 ARSEDSRPAAHFHLSSRRRHQSGMGVHGDYVIGNDNERNYSY-QGHFOTRD 313

Db 426 EDEVSALGCHQAPSHRQSHQDEETGHGQR--GSIKEMSHHPGHTTVVKD 473

RESULT 8

S03170

homeotic protein cut - fruit fly (Drosophila melanogaster)

C:Species: Drosophila melanogaster

C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999

C:Accession: S03170

R:Blöchliger, K.; Bodmer, R.; Jack, J.; Jan, L.Y.; Jan, Y.N.

Nature 333, 629-635, 1988

A:Title: Primary structure and expression of a product from cut, a locus involved in spe

A:Reference number: S03170; MUID:88232956; PMID:2897632

A:Accession: S03170

A:Molecule type: mRNA

A:Residues: 1-2175 <BLO>

A:Cross-references: EMBL:X07985; NID:g7767; PIDN:CAA30794.1; PID:g7768

C:Genetics:

A:Gene: cut

A:Cross-references: FlyBase:FBgn0004198

C:Superfamily: homeotic protein cut; cut repeat homology; homeobox homology

C:Keywords: DNA binding; homeobox; nucleus; transcription regulation

F:896-958/Domain: cut repeat homology <CU1>

F:1339-1411/Domain: cut repeat homology <CU2>

F:1617-1689/Domain: cut repeat homology <CU3>

F:1746-1802/Domain: homeobox homology <HOX>

Query Match 6.3%; Score 137; DB 1; Length 2175;

Best Local Similarity 21.8%; Pred. No. 1.1;

Matches 64; Conservative 45; Mismatches 111; Indels 74; Gaps 9;

Qy 21 PAKATSTATQRRTRQLPLVLGFIGLVAVAILALTIWQTIVSHLDKELKSLKRVVDN 80









GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 13, 2004, 10:23:02 ; Search time 12 Seconds  
(without alignments)  
1774.724 Million cell updates/sec

Title: US-09-813-329-6  
Perfect score: 2162  
Sequence: 1 MTATLKPFTPTSANDGF.....DRNAVLREGNNRSYFGIFKV 409

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_42:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	153.5	7.1	852	1 SRCH_RABIT	P16230 oryctolagus
2	137	6.3	699	1 SRCH_HUMAN	P23327 homo sapien
3	137	6.3	2175	1 HMCU_DROME	P10180 drosophila
4	134	6.2	466	1 SRO9_YEAST	P25567 saccharomyc
5	132	6.1	391	1 EDA_BOVIN	Q9beg5 bos taurus
6	132	6.1	391	1 EDA_HUMAN	Q92838 homo sapien
7	129	6.0	1085	1 IFH1_YEAST	P39520 saccharomyc
8	128	5.9	391	1 EDA_MOUSE	O54693 mus musculus
9	123.5	5.7	278	1 OSTP_SHEEP	Q9x8y9 ovis aries
10	123.5	5.7	332	1 Z265_RAT	O35986 rattus norv
11	122.5	5.7	713	1 NUC1_MESAU	P08199 mesocricetu
12	122.5	5.7	1379	1 TRP7_YEAST	P40340 saccharomyc
13	121.5	5.6	314	1 OSTP_HUMAN	P10451 homo sapien
14	121.5	5.6	326	1 Z265_MOUSE	Q9r020 mus musculus
15	121	5.6	337	1 Z265_HUMAN	O95218 homo sapien
16	120.5	5.6	1042	1 SUWA_DROME	P12297 drosophila
17	120.5	5.6	1161	1 NRDC_MOUSE	O8bhg1 mus musculus
18	120	5.6	294	1 OSTP_MOUSE	P10923 mus musculus
19	119.5	5.5	303	1 OSTP_PIG	P14287 sus scrofa
20	118.5	5.5	1105	1 A3B1_MOUSE	Q9z1t1 mus musculus
21	118.5	5.5	4835	1 MDN1_GIALA	O8c5t1 giardia lam
22	118	5.5	514	1 TUBE_DROVI	Q08171 drosophila
23	118	5.5	1067	1 SGG_DROME	P18431 drosophila
24	117.5	5.4	406	1 SR40_YEAST	P32583 saccharomyc
25	117	5.4	657	1 YGG6_YEAST	P53165 saccharomyc
26	117	5.4	1014	1 HEX2_YEAST	Q00816 saccharomyc
27	116	5.4	249	1 TN12_HUMAN	O43508 homo sapien
28	116	5.4	1025	1 MK21_YEAST	Q12176 saccharomyc
29	116	5.4	4910	1 MDN1_YEAST	Q12019 saccharomyc
30	115.5	5.3	706	1 NUC1_MOUSE	P09405 mus musculus
31	115.5	5.3	1083	1 HOS4_YEAST	P40480 saccharomyc
32	115	5.3	590	1 YNW7_YEAST	P53863 saccharomyc
33	114.5	5.3	414	1 NSR1_YEAST	P27476 saccharomyc

34	114.5	5.3	441	1 YB6E_YEAST	P38193 saccharomyc
35	114.5	5.3	802	1 NAB3_YEAST	P38996 saccharomyc
36	114.5	5.3	1435	1 EBA1_PLAFC	P19214 plasmodium
37	114	5.3	899	1 YMJ3_YEAST	O04500 saccharomyc
38	113.5	5.2	277	1 OSTK_BOVIN	P31098 bos taurus
39	113.5	5.2	851	1 BCK2_YEAST	P33306 saccharomyc
40	113.5	5.2	1807	1 VTA2_XENLA	P18709 xenopus lae
41	113	5.2	278	1 OSTP_BOVIN	P31096 bos taurus
42	113	5.2	510	1 GAT1_YEAST	P43574 saccharomyc
43	113	5.2	646	1 SGI_BOVIN	P23389 bos taurus
44	112.5	5.2	553	1 WRK6_ARATH	Q9C519 arabidopsis
45	112.5	5.2	1253	1 DSPP_HUMAN	Q9nzw4 homo sapien

ALIGNMENTS

RESULT 1  
SRCH\_RABIT  
ID SRCH\_RABIT STANDARD; PRT; 852 AA.  
AC P16230;  
DT 01-APR-1990 (Rel. 14, Created)  
DT 01-APR-1990 (Rel. 14, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Sarcoplasmic reticulum histidine-rich calcium-binding protein precursor.  
DE HRC OR HCP.  
GN Oryctolagus cuniculus (Rabbit).  
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
OX NCBI\_TaxID=9986;  
RN [1]  
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
RC TISSUE=Skeletal muscle.  
RX MEDLINE=9003684; PubMed=2808365;  
RA Hofmann S.L., Goldstein J.L., Orth K., Moomaw C.R., Slaughter C.A., Brown M.S.;  
RA "Molecular cloning of a histidine-rich Ca2+-binding protein of sarcolemmal reticulum that contains highly conserved repeated elements";  
RL J. Biol. Chem. 264:18083-18090(1989).  
CC -!- FUNCTION: May play a role in the regulation of calcium sequestration or release in the SR of skeletal and cardiac muscle.  
CC -!- SUBCELLULAR LOCATION: Sarcoplasmic reticulum lumen.  
CC -!- SIMILARITY: STRONG, TO HUMAN HRC.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; J05080; AAA31279.1; -;  
CC PIR; A34373; A34373.  
CC InterPro; IPR002134; HCP.  
CC PROSITE; PS00328; HCP; 10.  
KW Calcium-binding; Signal; Repeat.  
FT SIGNAL 1 27  
FT CHAIN 28 852  
FT -----  
FT MOD RES 28 28  
FT DOMAIN 59 100  
FT REPEAT 59 79  
FT REPEAT 80 100  
FT REPEAT 1-1.  
FT REPEAT 1-2.  
FT DOMAIN 199 470  
FT REPEAT 224 224  
FT REPEAT 2-1.  
FT REPEAT 225 253  
FT REPEAT 2-2.  
FT REPEAT 254 282  
FT REPEAT 2-3.  
FT REPEAT 283 310  
FT REPEAT 2-4.  
FT REPEAT 311 339  
FT REPEAT 2-5.  
FT REPEAT 340 367  
FT REPEAT 2-6.

POTENTIAL.  
SARCOPLASMIC RETICULUM HISTIDINE-RICH  
CALCIUM-BINDING PROTEIN.  
BLOCKED.  
2 X APPROXIMATE TANDEM REPEATS.  
1-1.  
10 X TANDEM REPEATS, ACIDIC.



RX MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galie R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe W., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfankuch C., Baldwin D.,  
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.V., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,  
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davidson L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,  
 RA Foeller C., Gabriellista A.E., Garg N.S., Gelbart W.M., Glaeser K.,  
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.B., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C.,  
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Laeko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milahina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskaas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of *Drosophila melanogaster*.";  
 RL Science 287:2185-2195 (2000).  
 CC -1- FUNCTION: Regulator of cell fate decisions in multiple lineages.  
 CC Specifically, functions as a determination factor that specifies  
 CC sensory organ identity in precursor cells. Probably also involved  
 CC in cell type specification of Malpighian tubules. In absence of  
 CC cut gene external sensory organs are transformed into chordotonal  
 CC organs.  
 CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).  
 CC -1- TISSUE SPECIFICITY: Detected in many cells in the central nervous  
 CC system, all external sensory organs, some peripheral neurons, and  
 CC in the non-neural cells of the spiracles and the Malpighian  
 CC tubules.  
 CC -1- DEVELOPMENTAL STAGE: Cell-specific pattern of expression. Broadly  
 CC expressed during embryonic development.  
 CC -1- DOMAIN: Asn at position 47 of the homeobox may participate in  
 CC regulating DNA-binding activity by promoting homo- and  
 CC heterodimerization.  
 CC -1- SIMILARITY: Belongs to the CUT homeobox family.  
 CC -1- SIMILARITY: Contains 1 homeobox domain.  
 CC -1- SIMILARITY: Contains 3 CUT domains.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; X07985; CAA30794.1; -;  
 CC EMBL; AE003441; AAP46264.2; -;  
 CC PIR; S03170; S03170.  
 CC TRANSFAC; T02004; -;  
 CC FlyBase; FBgn0004196; ct.  
 CC GO; GO:0005634; C:nucleus; IDA.

DR GO; GO:0005857; P:wing margin morphogenesis; NAS.  
 DR InterPro; IPR007108; Cut homeo.  
 DR InterPro; IPR003350; Hmoec CUT.  
 DR InterPro; IPR001356; Homeobox.  
 DR Pfam; PF02376; CUT; 3.  
 DR Pfam; PF00046; homeobox; 1.  
 DR ProDom; PD000010; Homeobox; 1.  
 DR SMART; SM00389; HOX; 1.  
 DR PROSITE; PS00027; HOMEBOX 1; 1.  
 DR PROSITE; PS00071; HOMEBOX 2; 1.  
 DR Transcription regulation; Homeobox; DNA-binding; Coiled coil.  
 DR Developmental protein; Nuclear protein; Repeat; Coiled coil.  
 FT DOMAIN 265 343 COILED COIL (POTENTIAL).  
 FT DOMAIN 433 499 COILED COIL (POTENTIAL).  
 FT DNA BIND 798 964 CUT 1.  
 FT DOMAIN 1056 1161 COILED COIL (POTENTIAL).  
 FT DNA BIND 1329 1417 CUT 2.  
 FT DOMAIN 1463 1522 COILED COIL (POTENTIAL).  
 FT DNA BIND 1608 1695 CUT 3.  
 FT DNA BIND 1745 1804 HOMEBOX.  
 FT DOMAIN 194 210 ALA/GLN-RICH.  
 FT DOMAIN 235 243 ALA-RICH.  
 FT DOMAIN 271 293 ASP/GLU-RICH (ACIDIC).  
 FT DOMAIN 384 428 ASN-RICH.  
 FT DOMAIN 547 554 ASP/GLU-RICH (ACIDIC).  
 FT DOMAIN 574 584 ASP/GLU-RICH (ACIDIC).  
 FT DOMAIN 616 630 ALA-RICH.  
 FT DOMAIN 665 699 HIS/GLN-RICH (OPA-REPEAT).  
 FT DOMAIN 2004 2014 ALA-RICH.  
 FT DOMAIN 2071 2077 ASP/GLU-RICH (ACIDIC).  
 FT DOMAIN 2124 2136 ALA/PRO-RICH.  
 SQ SEQUENCE 2175 AA; 233628 MW; 08BF80C4861BD0AB CRC64;  
 Query Match 6.3%; Score 137; DB 1; Length 2175;  
 Best Local Similarity 21.8%; Pred. No. 0.6;  
 Matches 64; Conservative 45; Mismatches 11; Indels 74; Gaps 9;  
 Qy 21 PAKATSTATQRRTRQLPLVLGFLGLVAILALTIWQTRVSHLDKELSLKRVVDN 80  
 Db 227 PAATVATGAAAAAATPIATGNVSGSTTSNANH-NSNSHQDE----- 273  
 Qy 81 LQORLGINYLDEDFEFOKEVENALIDYPKKVGLTDEEDDDGDLSDADEDDVSYS 140  
 Db 274 -----ELDDEEDEDEDE-----DEDEENASMQSNADMDLDAQQE 311  
 Qy 141 SVDDVGR-----DYEDYD-----MLANKLNAHTGTTPTSETTAEGEGETDS-- 182  
 Db 312 TRTEPSATTOOQHQOQDTELEENKQAGEASLVSNHNHTDSNNCSKRNNNGNSESQ 371  
 Qy 183 --ASSASNDNVDFDDFTSYNAHKKQERKSRIADVNEBQNIQGNHTELQEKSSNE-- 237  
 Db 372 HVASSAEDDDCANNTNTSNNNTSNTATSTNT--NNNNNNSSGSEKRRKNNNNNG 429  
 Qy 238 -----ATSKESPAPLHRRHRMHSRHHLLVKARSEDSRPAAHFLSSRRRH 284  
 Db 430 QPAVLLAAKDKETKALLDELQRLRAQOQTHLVQIRLEE-----HLEVKRQH 476  
 RESULT 4  
 SRO9 YEAST STANDARD; PRT; 466 AA.  
 AC P25567;  
 DT 01-MAY-1992 (Rel. 22, Created)  
 DT 01-MAY-1992 (Rel. 22, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE SRO9 protein.  
 GN SRO9 OR YCL037C OR YCL37C.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
 OK NCBI\_TaxID=4932;  
 RN [1]  
 RP SEQUENCE FROM N.A.







```

CC CC Name=I;
CC CC IsoId=Q92838-2; Sequence=VSP_006454, VSP_006455;
CC CC Name=A2;
CC CC IsoId=Q92838-3; Sequence=VSP_006464;
CC CC Name=B;
CC CC IsoId=Q92838-4; Sequence=VSP_006462, VSP_006463;
CC CC Name=C;
CC CC IsoId=Q92838-5; Sequence=VSP_006458, VSP_006461;
CC CC Name=D;
CC CC IsoId=Q92838-6; Sequence=VSP_006456, VSP_006457;
CC CC Name=E;
CC CC IsoId=Q92838-7; Sequence=VSP_006459, VSP_006461;
CC CC Name=F;
CC CC IsoId=Q92838-8; Sequence=VSP_006460, VSP_006461;
CC CC TISSUE SPECIFICITY: Not abundant; expressed in specific cell types
CC CC of ectodermal (but not mesodermal) origin of keratinocytes, hair
CC CC follicles, sweat glands. Also in adult heart, liver, muscle,
CC CC pancreas, prostate, fetal liver, uterus, small intestine and
CC CC umbilical chord.
CC CC -!- PTM: N-glycosylated.
CC CC -!- PTM: Processing by furin produces a secreted form.
CC CC -!- DISEASE: Defects in ED1 are the cause of ectodermal dysplasia,
CC CC anhidrotic (EDA) [MIM:305100]; also known as X-linked hypohidrotic
CC CC ectodermal dysplasia (XUHED). EDA is a disease characterized by
CC CC sparse hair (atrichosis or hypotrichosis), abnormal or missing
CC CC teeth and the inability to sweat due to the absence of sweat
CC CC glands. EDA is the most common form of over 150 clinically
CC CC distinct ectodermal dysplasias. This disease was already described
CC CC by Darwin.
CC CC -!- SIMILARITY: Belongs to the tumor necrosis factor family.
CC CC -!- SIMILARITY: Contains 1 collagenous domain.
CC CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC CC or send an email to license@isb-sib.ch).
CC CC -----
CC CC EMBL; U59227; AAC50678.1; -
CC CC EMBL; U59228; AAC50679.1; -
CC CC EMBL; AF061189; AAC77371.1; -
CC CC EMBL; AF061190; AAC77372.1; -
CC CC EMBL; AF061191; AAC77373.1; -
CC CC EMBL; AF061192; AAC77374.1; -
CC CC EMBL; AF061193; AAC77375.1; -
CC CC EMBL; AF061194; AAC77376.1; -
CC CC EMBL; AF060999; AAC36302.1; -
CC CC EMBL; AF060998; AAC36303.1; -
CC CC EMBL; AF060992; AAC36303.1; JOINED.
CC CC EMBL; AF060993; AAC36303.1; JOINED.
CC CC EMBL; AF060994; AAC36303.1; JOINED.
CC CC EMBL; AF060995; AAC36303.1; JOINED.
CC CC EMBL; AF060996; AAC36303.1; JOINED.
CC CC EMBL; AF060997; AAC36303.1; JOINED.
CC CC EMBL; AF040628; AAC77363.1; -
CC CC EMBL; AL158069; CAD18890.1; -
CC CC EMBL; AL158141; CAD13493.1; -
CC CC Genew; HGNC:3157; ED1.
CC CC MIM; 300451; -
CC CC MIM; 305100; -
CC CC GO; GO:0005856; C:cytoskeleton; TAS.
CC CC GO; GO:0016021; C:integral to membrane; TAS.
CC CC GO; GO:0005624; C:membrane fraction; TAS.
CC CC GO; GO:0005886; C:plasma membrane; TAS.
CC CC GO; GO:0005102; F:receptor binding; TAS.
CC CC GO; GO:0007398; P:ectoderm development; TAS.
CC CC GO; GO:0007165; P:signal transduction; TAS.
CC CC InterPro; IPR008160; Collagen.
CC CC InterPro; IPR006052; TNF family.
CC CC InterPro; IPR008983; TNF_like.
CC CC -----
Query Match 6.1%; Score 132; DB 1; Length 391;
Best Local Similarity 28.6%; Pred.No. 0.17;
Matches 44; Conservative 31; Mismatches 63; Indels 16; Gaps 8;
Oy 264 KARSERSRPAHPHLLSSRRRHQGS-MGYHCDMYIG--NDNER---NSYQGHFQTRGVLT 317
Db 240 KAGTRENQPAV-VHLQG---QGSALQVKNLDSGGVLDNWSRITMNPVKFKLHPRSGELE 294
Oy 318 VTNTGLYYVYAQ--ICVNNSHDQNGFIVFOGDPFLQCLNTVPTNMPHKVHTCHTSLIH 375
Db 295 VLVDGTGYFYISQVEVYVYINFTDFASVEVVVDEKPFLOCTRSIETGKTN-YNICYTAGVCL 353
Oy 376 LERNERHLKDHNDRNAVLRGNRSYFGIFKV 409
Db 354 LKARQKIAVMVHADIS--INMSKHTTFFGAIRL 385
RESULT 7
IFHL YEAST STANDARD; PRT; 1085 AA.
AC P39520;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE IFH1 protein (RRP3 protein)
GN IFH1 OR RRP3 OR YLR223C OR L8083.9.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=ATCC 28383 / FL100;
RX MEDLINE=95304839; PubMed=7785326;
RA Cheral I., Thuriaux P.;
RT "The IFH1 gene product interacts with a fork head protein in
RL Yeast 11:261-270(1995).
RN [2]
SEQUENCE FROM N.A.
RC STRAIN=S288c / AB972;
RX MEDLINE=97313267; PubMed=9169871;
RA Johnston M., Hillier L., Riles L., Albertmann K., Andre B., Ansorge W.,
RA Benes V., Brueckner M., Delius H., Dubois E., Duesterhoeft A.,
RA Entian K.-D., Floeth M., Goffeau A., Hebling U., Heumann K.,
RA Heuss-Neitzel D., Hilbert H., Hilger F., Kleine K., Koetter P.,
RA Louis E.-J., Messenguy F., Mewes H.-W., Miosga T., Moestl D.,
RA Mueller-Auer S., Nentwich U., Obermaier B., Piravandi E., Pohl T.M.,
RA Portetelle D., Purnelle B., Rechmann S., Rieger M., Rinke M., Rose M.,
RA Scharfe M., Scherens B., Scholler P., Schwager C., Schwarz S.,
RA Underwood A.P., Urrestarazu L.A., Vandenbol M., Verhasselt P.,
RA Vierendeels F., Voet M., Volckaert G., Voss H., Wambutt R., Wedler E.,
RA Wedler H., Zimmermann F.K., Zollner A., Hani J., Hoheisel J.D.;
RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome XII.";
RL Nature 387:87-90(1997).
CC -!- FUNCTION: Controls the pre-rRNA processing machinery in
CC conjunction with FHL1. Could convert FHL1 from a repressor
CC to an activator.
CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC CC -----
CC CC EMBL; U29488; CAA82624.1; -
CC CC EMBL; U19027; AAB67412.1; -
CC CC PIR; S55352; S55352.
CC CC GenOnline; 142285; -.
CC CC SGD; S0004213; IFH1.

```



DR GO; GO:0006348; P:chromatin silencing at telomere; IMP.  
DR GO; GO:0006364; P:rRNA processing; IGI.  
KW Nuclear protein; Transcription regulation.  
FT DOMAIN 122 163 ASP/GLU-RICH (HIGHLY ACIDIC).  
SQ SEQUENCE 1085 AA; 122491 MW; 122491 MW; BE1C7DF06213FE0 CRC64;  
Query Match 6.0%; Score 129; DB 1; Length 1085;  
Best Local Similarity 19.9%; Pred. NO. 0.85;  
Matches 73; Conservative 48; Mismatches 126; Indels 120; Gaps 12;  
QY 69 KELSLKRVNQLQRLGIVYLDFEDFQ-----KEYENALIDYPKVGDGLTDEEDDDG 123  
DB 94 KKSLLIQRIQINDDEG-----TESSDYQAVTDGESENEBEESEEDDDDDDDDD 147  
QY 124 DGLSDIADDDDDSVSYVDVGDYED-----YTDMLNKLN 160  
DB 148 DGSDDSDSE-----TSDDENIDFVKLTAQRKRAMKALSAMNTNNTLYSRENSNK 201  
QY 161 NAHTGTTTPTSETTAB-----GEGETDSASSANDNDNVDFDF 196  
DB 202 NKSVKLSPKNEEBEQKEKEKEKEEQQESNKKEVNGSGTTTQQAQSFKFKKEDDG 261  
QY 197 TSY-NAHKKKOERKSRSTADYVNEQN-----IQGNHYELOKSNEATSKESP 244  
DB 262 ISFGNGEGYNEDIGEEVLDLKNNNGNEBKLDKSKYMLGNDELRFNISESDESE-- 319  
QY 245 APLHRRMRHSHRHLVRKARSDRPAAHFHLSSRRRHQSGWY-----HGDWYIGND 299  
DB 320 -----YDIDQDAFDVINNEDSGEIGTDLTGEDDLPLEE 356  
QY 300 NERN-----SYQG--HFQTRDGLVTVNTGLVYVYQA-ICYNNSHDQNGFVIFQ 345  
DB 357 EEQNIIVSELQNDDELSPDGSFTHIEGSDPDAENKFLQNEYNGVDEEDDEDEMSD 416  
QY 346 GDTFELQ 352  
DB 417 FDMPPFE 423  
RESULT 8  
EDA\_MOUSE  
ID EDA\_MOUSE STANDARD; PRT; 391 AA.  
AC 054693; 035705; Q9QWJ8; Q9QZ01; Q9QZ02;  
DT 15-JUL-1999 (Rel. 38, Created)  
DT 15-JUL-1999 (Rel. 38, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Ectodysplasin A (EDA protein homolog) (Tabby protein).  
GN ED1 OR EDA OR TA.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
[1]  
RN SEQUENCE FROM N.A. (ISOFORMS TAA; TAB AND TAC).  
RC STRAIN=129/SV;  
RX MEDLINE=98058770; PubMed=9371801;  
RA Srivastava A.K., Piepa J., Hartung A.J., Du Y., Ezer S., Jenks T.,  
RA Shimada T., Pekkanen M., Mikkola M.L., Ko M.S.H., Thesleff I.,  
RA Kere J., Schlessinger D.;  
RT "The tabby phenotype is caused by mutation in a mouse homologue of the  
RT EDA gene that reveals novel mouse and human exons and encodes a  
RT protein (ectodysplasin-A) with collagenous domains.";  
RL Proc. Natl. Acad. Sci. U.S.A. 94:13069-13074(1997).  
[2]  
RN SEQUENCE FROM N.A. (ISOFORM TAD).  
RX MEDLINE=9749184; PubMed=9285798;  
RA Ferguson B.M., Brockdorff N., Formstone E., Nguyen T.,  
RA Kronmiller J.E., Zonana J.;  
RT "Cloning of Tabby, the murine homolog of the human EDA gene: evidence  
RT for a membrane-associated protein with a short collagenous domain.";  
RL Hum. Mol. Genet. 6:1589-1594(1997).  
[3]  
RN SEQUENCE FROM N.A. (ISOFORMS TA-A2 AND TA-A3).  
RP

RC TISSUE=Embryo;  
RX MEDLINE=20005791; PubMed=10534613;  
RA Mikkola M.L., Piepa J., Pekkanen M., Paulin L., Nieminen P., Kere J.,  
RA Thesleff I.;  
RT "Ectodysplasin, a protein required for epithelial morphogenesis, is a  
RT novel TNF homologue and promotes cell-matrix adhesion.";  
RL Mech. Dev. 88:133-146(1999).  
CC -!- FUNCTION: Involved in epithelial-mesenchymal signaling during the  
CC morphogenesis of ectodermal organs. Isoform TAA binds only to the  
CC receptor EDAR, while isoform TA-A2 binds exclusively to the  
CC receptor XEDAR.  
CC -!- SUBUNIT: Homotrimer (By similarity).  
CC -!- SUBCELLULAR LOCATION: Type II membrane protein and secreted (By  
CC similarity).  
CC -!- ALTERNATIVE PRODUCTS:  
CC Event=Alternative splicing; Named isoforms=6;  
CC Comment=Additional isoforms seem to exist;  
CC Name=TAA; Synonyms=A1;  
CC IsoId=O54693-1; Sequence=Displayed;  
CC Name=TA-A2;  
CC IsoId=O54693-2; Sequence=VSP\_006471;  
CC Name=TA-A3;  
CC IsoId=O54693-3; Sequence=VSP\_006469, VSP\_006471;  
CC Name=TAB;  
CC IsoId=O54693-4; Sequence=VSP\_006466, VSP\_006467;  
CC Name=TAC;  
CC IsoId=O54693-5; Sequence=VSP\_006465, VSP\_006468;  
CC Name=TAD;  
CC IsoId=O54693-6; Sequence=VSP\_006470;  
CC -!- PTM: N-glycosylated (By similarity).  
CC -!- PTM: Processing by furin produces a secreted form (By similarity).  
CC -!- DISEASE: Defects in ED1 are the cause of the tabby phenotype in  
CC mice (the equivalent of anhidrotic ectodermal dysplasia in  
CC humans). The disease is characterized by sparse hair (atricichosis  
CC or hypotrichosis), abnormal or missing teeth and the inability to  
CC sweat due to the absence of sweat glands.  
CC -!- SIMILARITY: Belongs to the tumor necrosis  
CC -!- SIMILARITY: Contains 1 collagenous domain.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; AF016627; AAB95202.1; -;  
CC EMBL; AF016628; AAB95203.1; -;  
CC EMBL; AF016629; AAB95204.1; -;  
CC EMBL; AF016630; AAB95205.1; -;  
CC EMBL; AF016631; AAB95206.1; -;  
CC EMBL; AF004434; AAB88121.1; -;  
CC EMBL; AF004435; AAB88122.1; -;  
CC EMBL; Y13438; CAA73849.1; -;  
CC EMBL; AJ243657; CAB52696.1; -;  
CC EMBL; AJ243658; CAB52697.1; -;  
CC MGD; MG11195272; Eda.  
CC GO; GO:0045177; C:apical part of cell; IDA.  
CC GO; GO:0005789; C:endoplasmic reticulum membrane; IDA.  
CC GO; GO:0005887; C:integral to plasma membrane; IDA.  
CC GO; GO:0007160; P:cell-matrix adhesion; IDA.  
CC GO; GO:0042346; P:positive regulation of NF-kappaB protein-nu. . . ; IDA.  
CC GO; GO:0007431; P:salivary gland development; IDA.  
CC InterPro; IPR008160; Collagen.  
CC InterPro; IPR006052; TNF family.  
CC InterPro; IPR008983; TNF like.  
CC Pfam; PF01391; Collagen; 1.  
CC SMART; SM00207; TNF; 1.  
CC PROSITE; PS00251; TNF\_1; FALSE\_NEG.  
CC PROSITE; PS00049; TNF\_2; 1.  
KW Differentiation; Developmental protein; Collagen; Transmembrane;  
KW Signal-anchor; Glycoprotein; Alternative splicing.





```
FT MOD_RES 665 665 METHYLATION (DI-).
FT MOD_RES 669 669 METHYLATION (DI-).
FT MOD_RES 673 673 METHYLATION (DI-).
FT MOD_RES 679 679 METHYLATION (DI-).
FT MOD_RES 681 681 METHYLATION (DI-).
FT MOD_RES 687 687 METHYLATION (DI-).
FT MOD_RES 691 691 METHYLATION (DI-).
FT MOD_RES 694 694 METHYLATION (DI-).
SQ SEQUENCE 713 AA; 76997 MW; 79DDCF724CED7DB4 CRC64;

Query Match 5.7%; Score 122.5; DB 1; Length 713;
Best Local Similarity 20.0%; Pred. No. 1.3;
Matches 53; Conservative 40; Mismatches 109; Indels 63; Gaps 8;

QY 98 KEYENALIDYPKKVGDLTDEEDDDGGDLSTADDEDDVSYSSVDDVGADYEDYTDMLN 157
Db 133 KNGKNA-----KKDSDEDDDDDDDDDDDEDEDEED-----EFEP----- 170
QY 158 KLNNHAHTGTTPTSETTAEGETDSASSASNDNDVDFDFTSYNA-----HKKKQER-- 208
Db 171 PVVKGQKGVAAAPASEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDE 230
QY 209 ---KRSIADVNEEQNIQGNHTLOEKSNEATSK-----SPAPLHRRRMRHRRH 259
Db 231 PVKAKNAEEDDDDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDE 290
QY 260 LLVRKARSEDSPAAHFLSLRRRHQSGMGVHGYDNYGNDNERN-----YQHFRQTR 312
Db 291 PEAKQKVEGESITPPFNL-----FIGNLPNKSVAELKVAISEPFAKN 334
QY 313 DGLVTVTNTGLYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVY 337
Db 335 DLAUVVDVRTGTRKRGYVDFESAED 359

RESULT 12
TBP7_YEAST
ID TBP7_YEAST STANDARD; PRT; 1379 AA.
AC P40340;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE TAT-binding homolog 7.
GN YTA7 OR YGR270W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288c / FY1679;
RA Agostoni Carbone M.L.; Lucchini G.; Melchiorretto P.; Nardese V.;
RA Vanoni M.; Panzeri L.;
RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=S288c;
RA Feldmann H.;
RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RP PRELIMINARY SEQUENCE OF 1-1207 FROM N.A.
RC STRAIN=S288c;
RX MEDLINE=95274317; PubMed=7754704;
RA Schnall R.; Mannhaupt G.; Stucka R.; Tauer R.; Ehnle S.;
RA Schwarzlose C.; Vetter I.; Feldmann H.;
RT "Identification of a set of yeast genes coding for a novel family of
RT putative ATPases with high similarity to constituents of the 26S
RT protease complex.";
RL Yeast 10:1141-1155(1994).
CC -1- SIMILARITY: Belongs to the AAA ATPase family.
CC -1- SIMILARITY: Contains 1 bromodomain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
```

```
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Y07893; CAA69201.1; -
DR EMBL; Z73055; CAA97300.1; -
DR EMBL; X81072; CAA56963.1; -
DR PIR; S64603; S64603.
DR GERMOnline; 141582; -
DR SGD; S0003502; YTA7.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR003959; AAA_ATPase_cent.
DR InterPro; IPR003960; AAA_sub.
DR InterPro; IPR001487; Bromodomain.
DR Pfam; PF00004; AAA; 1.
DR Pfam; PF00439; bromodomain; 1.
DR SMART; SM00382; AAA; 1.
DR PROSITE; PS00674; AAA; 1.
DR PROSITE; PS00633; BROMODOMAIN_1; FALSE_NEG.
DR PROSITE; PS00014; BROMODOMAIN_2; 1.
KW ATP-binding; Bromodomain.
FT DOMAIN 135 139 POLY-GLU.
FT DOMAIN 157 160 POLY-ARG.
FT DOMAIN 358 361 POLY-ASN.
FT DOMAIN 390 395 POLY-LYS.
FT NP_BIND 454 461 ATP (POTENTIAL).
FT DOMAIN 737 740 POLY-GLU.
FT DOMAIN 1044 1086 BROMODOMAIN (DIVERGENT).
FT CONFLICT 70 70 D -> E (IN REF. 2).
FT CONFLICT 241 241 S -> N (IN REF. 2).
FT CONFLICT 1016 1016 S -> N (IN REF. 2).
FT CONFLICT 1142 1142 S -> N (IN REF. 2).
FT CONFLICT 1153 1153 K -> E (IN REF. 2).
FT CONFLICT 1276 1276 I -> R (IN REF. 2).
FT CONFLICT 1283 1283 Q -> P (IN REF. 2).
SQ SEQUENCE 1379 AA; 157406 MW; 31D1F6F87E62E04F CRC64;

Query Match 5.7%; Score 122.5; DB 1; Length 1379;
Best Local Similarity 23.8%; Pred. No. 2.9;
Matches 53; Conservative 40; Mismatches 101; Indels 29; Gaps 9;

QY 87 INVLDEFEFOKEYENALIDYPK-KVQGLTDEEDDDGGDLSTADDEDDVSYSSVDDV 145
Db 44 INYAEIKVDFLEDDQVMDKDETPTVTSDEHHNNQK-----DDEDDVDLVSPHEN 98
QY 146 GADYEDYTDMLN-KLNNHAHTGTTPTSETTAEGETDSASSASNDNDVDFDFTSYNAHK 204
Db 99 ARTNEELTNERMLRKKAH---DPEEDDESFEEDVDDDEEEEADEFEDYLEDSDKN 155
QY 205 KQERSRS-----IADVNEEQNIQGNHTLOEKSNEATSKESAPLHRRRMRHRRHL 260
Db 156 NRRRAADRKVFVDPDDDEYDEDD--EGDRISHSASSK-----RLKANSRR--- 203
QY 261 LVRKARSEDSPAAHFLSLRRRHQSGMGVHGYDNYGNDNERN 303
Db 204 -TRSRHPETPPVRRRAURSRHSRTSNEEND--DENDNSRN 243

RESULT 13
OSTP_HUMAN
ID OSTP_HUMAN STANDARD; PRT; 314 AA.
AC P10451; Q15681; Q15682; Q15683;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Osteopontin precursor (Bone sialoprotein 1) (Urinary stone protein)
DE (Secreted phosphoprotein 1) (SPP-1) (Nephropontin) (Uropontin).
OS SPPI OR OPN.
GN Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```

OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89263749; PubMed=2726470;
RA Kiefer M.C., Bauer D.M., Barr P.J.;
RT "The cDNA and derived amino acid sequence for human osteopontin.";
RL Nucleic Acids Res. 17:3306-3306(1989).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=90353945; PubMed=1974876;
RA Young M.F., Kerr J.M., Termini J.D., Wewer U.M., Wang M.G.,
RA McBride O.W., Fisher L.W.;
RT "cDNA cloning, mRNA distribution and heterogeneity, chromosomal
RT location, and RFLP analysis of human osteopontin (OPN).";
RL Genomics 7:491-502(1990).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=92108068; PubMed=1729712;
RA Shiraga H., Min W., Vandusen W.J., Clayman M.D., Miner D.,
RA Terrell C.H., Sherbotie J.R., Foreman J.W., Przysiecki C.,
RA Neilson E.G., Hoyer J.R.;
RT "Inhibition of calcium oxalate crystal growth in vitro by uropontin:
RT another member of the aspartic acid-rich protein superfamily.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:426-430(1992).
RN [4]
RP SEQUENCE FROM N.A.
RA Crosby A.H., Edwards S., Murray J.C., Dixon M.J.;
RL Submitted (FEB-1995) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=95031968; PubMed=7945249;
RA Hijiya N., Setoguchi M., Matsuura K., Higuchi Y., Akizuki S.,
RA Yamamoto S.;
RT "Cloning and characterization of the human osteopontin gene and its
RT promoter.";
RL Biochem. J. 303:255-262(1994).
RN [6]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Yu W., Sarginson J., Gibbs R.A.;
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
RN [7]
RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RX MEDLINE=95139605; PubMed=7637791;
RA Saitoh Y., Kuratsu J., Takeshima H., Yamamoto S., Ushio Y.;
RT "Expression of osteopontin in human glioma. Its correlation with the
RT malignancy.";
RL Lab. Invest. 72:55-63(1995).
RN [8]
RP SEQUENCE FROM N.A. (ISOFORM A).
RC TISSUE=Kidney;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max J., Wang J., Hesieh F.,
RA Diatchenko L., Maruina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udell T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raba S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A.C., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grinwald J., Schmutz J., Myers R.M.,
RA Butlerfield V.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
```

---

```

RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [9]
RP SEQUENCE OF 67-278 FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=92246977; PubMed=1575754;
RA Kohri K., Suzuki Y., Yoshida K., Yamamoto K., Amasaki N., Yamate T.,
RA Umekawa T., Iguchi M., Sinohara H., Kurita T.;
RT "Molecular cloning and sequencing of cDNA encoding urinary stone
RT protein, which is identical to osteopontin.";
RL Biochem. Biophys. Res. Commun. 184:859-864(1992).
CC -!- FUNCTION: Binds tightly to hydroxyapatite. Appears to form an
CC integral part of the mineralized matrix. Probably important to
CC cell-matrix interaction.
CC -!- FUNCTION: Acts as a cytokine involved in enhancing production of
CC interferon-gamma and interleukin-12 and reducing production of
CC interleukin-10 and is essential in the pathway that leads to type
CC I immunity (by similarity).
CC -!- SUBUNIT: Ligand for integrin alpha-V/beta-3.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=3;
CC Name=A; Synonym=OPN-A, OP1B;
CC IsoId=P10451-1; Sequence=Displayed;
CC Name=B; Synonym=OPN-B, OP1A;
CC IsoId=P10451-2; Sequence=VSP_003778;
CC Name=C; Synonym=OPN-C;
CC IsoId=P10451-3; Sequence=VSP_003777;
CC -!- PTM: Extensively phosphorylated on serine residues.
CC -!- PTM: N- and O-glycosylated.
CC -!- DISEASE: This protein plays a principal role in urinary stone
CC formation as the stone matrix.
CC -!- SIMILARITY: Belongs to the osteopontin family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X13694; CAA31984.1; -
CC EMBL; J04765; AAA59974.1; -
CC EMBL; M83248; AAA17675.1; -
CC EMBL; U20758; AAA68886.1; -
CC EMBL; AF052124; AAC28619.1; -
CC EMBL; D28759; BAA05949.1; -
CC EMBL; D28760; BAA05950.1; -
CC EMBL; D28761; BAA05951.1; -
CC EMBL; D14813; BAA03554.1; -
CC EMBL; BC017387; AAH17387.1; -
CC PIR; S50028; S09575.
CC GlycoSuiteDB; P10451; -.
CC Genew; HGNC:11255; SPPI.
CC MIM; 166490; -.
CC GO; GO:0005578; C:extracellular matrix; TAS.
CC GO; GO:0008189; F:apoptosis inhibitor activity; ISS.
CC GO; GO:0042056; F:chemotactant activity; TAS.
CC GO; GO:0005225; F:cytokine activity; ISS.
CC GO; GO:0003793; F:defense/immunity protein activity; TAS.
CC GO; GO:0008083; F:growth factor activity; TAS.
CC GO; GO:0005178; F:integrin binding; NAS.
CC GO; GO:0006916; P:anti-apoptosis; ISS.
CC GO; GO:0007267; P:cell-cell signaling; TAS.
CC GO; GO:0007160; P:cell-matrix adhesion; NAS.
CC GO; GO:0030595; P:immune cell chemotaxis; TAS.
CC GO; GO:0006954; P:inflammatory response; TAS.
CC GO; GO:0030502; P:negative regulation of bone mineralization; NAS.
CC GO; GO:0042102; P:positive regulation of T-cell proliferation; TAS.
CC GO; GO:0045637; P:regulation of myeloid blood cell differenti. .; TAS.
CC GO; GO:0042088; P:T-helper 1 type immune response; TAS.
CC InterPro; IPR002038; Osteopontin.
```

```
DR Pfam; PF00865; Osteopontin; 1.
DR PRINTS; PR00216; OSTEOPTNTN.
DR SMART; SM00017; OSTEO; 1.
DR PROSITE; PS00884; OSTEOPTNTN; 1.
KW Cytokine; Glycoprotein; Sialic acid; Biom mineralization; Cell adhesion;
KW Phosphorylation; Signal; Alternative splicing; Polymorphism.
FT SIGNAL 1 16 POTENTIAL.
FT CHAIN 17 314 OSTEOPTNTN.
FT SITE 159 161 CELL ATTACHMENT SITE.
FT CARBOHYD 79 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 106 106 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARSPLIC 31 57 Missing (in isoform C).
FT VARSPLIC 58 71 Missing (in isoform B).
FT VARSPLIC 301 301 /FTID=VSP 003778.
FT VARIANT 188 188 D -> H (in dbSNP:4660).
FT CONFLICT 237 237 D -> H (IN REF. 7).
FT CONFLICT 275 278 T -> A (IN REF. 7).
FT CONFLICT 278 SHEP -> GNSL (IN REF. 2).
SQ SEQUENCE 314 AA; 35422 MW; 4996429EC4752B86 CRC64;

Query Match 5.6%; Score 121.5; DB 1; Length 314;
Best Local Similarity 22.8%; Pred. No. 0.59;
Matches 61; Conservative 30; Mismatches 78; Indels 99; Gaps 14;

QY 91 DEFEFOKEYENALIDYPKK-----VDGLTDEDDDDGDLSDIADDDV-----S 138
D : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 64 EFTNFKQE-----TLPSKSNESHDMDDDDDDHVDSDQSDSDSDSDDDVDDTDDS 117

QY 139 YSSVDVGVADYED--YTDLNKL--NNAHTGTTPSETTARGEGET-----180
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 118 HQSDSHSHSDSDDELVTDFPTDLPAETVTPVPTVD--TYDGRGDSVYGLRSKKFRR 176

QY 181 -----DSASSANDNDVDPDFTSY 199
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 177 PDIQPDATDDEITSHMSEELNGAYKAIPVAQDLNAPSDWDSDGDSYETSQJDD-QSA 235

QY 200 NAAKKKQKRKRSIADVNEEQNIQGNHTELQKSSNEATSKESPAPLHRRHRMHSRRH 259
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 236 ETHSHKQSLYKKA---NDESN---EHSVDVDSQELSKVSR-----FHSHEPHS-HED 283

QY 260 LLV--RKARSDSRPAAHFHLSSRRRQ 285
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 284 MLVVDPKSKEEDK-----HLKFRISHE 305
```

```
RESULT 14
Z265 MOUSE
ID Z265 MOUSE STANDARD; PRT; 326 AA.
AC Q9R020;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Zinc finger protein 265 (zinc finger, splicing) (Fragment).
GN ZNF265 OR ZP265 OR Z18.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=20237530; PubMed=10773668;
RA Adams D.J., van der Weyden L., Kovacic A., Lovicu F.J., Copeland N.G.,
RA Gilbert D.J., Jenkins N.A., Ioannou P.A., Morris B.J.;
RT "Chromosome localization and characterization of the mouse and human
RT zinc finger protein 265 gene."
RL Cytogenet. Cell Genet. 88:68-73(2000).
CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -!- SIMILARITY: Contains 2 RANBP2-type zinc fingers.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
```

```
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF133818; AAF04474.1; -.
DR MGD; MGI:1858211; Zfp265.
DR InterPro; IPR001876; Znf-RanBP.
DR Pfam; PF00641; zf-RanBP; 2.
DR SMART; SM00547; ZNF_RBP2; 2.
DR PROSITE; PS01356; ZF_RANBP2_1; 2.
DR PROSITE; PS01356; ZF_RANBP2_2; 2.
KW Zinc-finger; Nuclear protein; Repeat.
FT ZN_FING 9 40 RANBP2-TYPE 1.
FT ZN_FING 65 94 RANBP2-TYPE 2.
FT DOMAIN 198 262 ARG/SER-RICH.
FT NON_TER 326 326
SQ SEQUENCE 326 AA; 36739 MW; 94593499E481FE88 CRC64;

Query Match 5.6%; Score 121.5; DB 1; Length 326;
Best Local Similarity 20.4%; Pred. No. 0.62;
Matches 61; Conservative 46; Mismatches 93; Indels 99; Gaps 12;

QY 20 FPAKATSTATAQRTRQLIPLVLGFTIGLVVA-----ILALTIWQTRVSHLDKELKSL 74
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 25 FARTSCNRCGREKTEAKMKAGGTEIGKTLAEKSGRLFSANDMQCKTCSNVNWARSE 84

QY 75 KRVD-----NLQRLG-----INYL-----EFDEFQKEYENALIDYPKVDG 113
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 85 CDMCTPKYAKLEERTGYGGGNRESVEYIERESDGEYDFGRKKK---YRGKAVG 140

QY 114 ---LTDEEDDDGDLSDIADDDVSVSSVDVGDYEDYDMLNKLNNHAHTGTTPTS 170
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 141 PASILKEVDKESGE---EDEDELSKYLD-----170

QY 171 ETARGEGETDSASSANDNDVDFDFTSYNAHKKQKRKRSIADVNEEQNIQGNHTEL 230
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 171 -----EDEDDADLSKYN---LDASEREDSKKKSNRRSR-----204

QY 231 QKSSNEATSKESPAPLHRRHRMHSRRHLLVRKARSDSRPAAHFHLSSRRRQSGMW 289
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 205 KRSRSHSRSSSRSSPSSSRSRSRSR-----SSSSSSSRSH---SGSRHSRSG 253
```

```
RESULT 15
Z265 HUMAN
ID Z265 HUMAN STANDARD; PRT; 337 AA.
AC Q95218; Q9UP63;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Zinc finger protein 265 (zinc finger, splicing).
GN ZNF265 OR Z18.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RC TISSUE=Fetal brain;
RX MEDLINE=99132651; PubMed=9931435;
RA Nakano M., Yoshitake K., Oikawa M., Miyoshi O., Yamada K., Kondo S.,
RA Miwa N., Soeda E., Jinno Y., Fujii T., Nishikawa N.;
RT "Identification, characterization and mapping of the human Z18 (zinc-
RT finger, splicing) gene."
RL Gene 225:59-65(1998).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM Z18-2).
RC TISSUE=Uterus;
RX MEDLINE=21154917; PubMed=11230166;
RA Wiemann S., Weill B., Wellenreuther R., Gassenhuber J., Glassl S.,
```

Job time : 14 secs

```
RA Anorge W., Boecker M., Bloecker H., Bauersachs S., Blum H.,
RA Lauber J., Duesterhoeft A., Beyer A., Koehrer K., Strack N.,
RA Meves H.-W., Ottenwaelder B., Obermaier B., Tampe J., Heubner D.,
RA Wambutt R., Korn B., Klein M., Poustka A.;
RT "Towards a catalog of human genes and proteins: sequencing and
RT analysis of 500 novel complete protein coding human cDNAs.";
RL Genome Res. 11:422-435(2001).
CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=ZIS-1;
CC IsoId=O95218-1; Sequence=Displayed;
CC Name=ZIS-2;
CC IsoId=O95218-2; Sequence=VSP 004493;
CC -!- SIMILARITY: Contains 2 RANBP2-type zinc fingers.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF065391; AAD09746.1; -
DR EMBL; AF065392; AAD09747.1; -
DR EMBL; AL136945; CAB66879.1; -
DR Genew; HGNC:13058; ZNF265.
DR MIM; 604347; -
DR GO; GO:0005634; C:nucleus; TAS.
DR GO; GO:0003723; F:RNA binding; TAS.
DR GO; GO:0003700; F:transcription factor activity; TAS.
DR GO; GO:0008380; P:RNA splicing; TAS.
DR InterPro; IPR001876; Znf_ranGDP.
DR Pfam; PF00641; zf-RanBP; 2.
DR SMART; SM00547; Znf_RBZ; 2.
DR PROSITE; PS01358; ZF_RANBP2_1; 2.
DR PROSITE; PS0199; ZF_RANBP2_2; 2.
DR Zinc-finger; Nuclear protein; Repeat; Alternative splicing.
DR ZN_FING 9 40 RANBP2-TYPE 1.
FT ZN_FING 65 94 RANBP2-TYPE 2.
FT DOMAIN 198 262 ARG/SER-RICH.
FT VARSPLOC 310 337 DATGHLDHPILVPVQKRNVLKFTS -> AR (in
FT isoform ZIS-2).
FT /FTId=VSP 004493.
SQ SEQUENCE 337 AA; 38223 MW; A4D20F6B3AD95925 CRC64;

Query Match 5.6%; Score 121; DB 1; Length 337;
Best Local Similarity 20.2%; Pred. No. 0.69;
Matches 60; Conservative 47; Mismatches 94; Indels 96; Gaps 11;

QY 20 FPAKATATATQRTQLIPLVLGFIGLGLVA-----ILALTIQTTRVSHLDKELKSL 74
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
25 FARTSCNCRGKREKTEAKMMKAGTEIGKTLAEKSGLFANDWQCKTCSNVNWARRE 84
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 75 KRVD-----NLQRLG-----INYL-----EFDEFOKEYENALIDYPKKVDG 113
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 85 CNMCTPKYAKLEERTGYGGFNERENVEYIEREESGEYDFGKKKK-----YRGKAVG 140
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 114 ---LTDEDDDDGGLSDIADEDDDDVSYSVDVGDYEDYDMLNKLNNHAHTGTTPTS 170
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 141 PASILKEVEDKESEGE-----EDEDELSKYKLD----- 170
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 171 ETTAEGGETDSASASNDNDVDFDFTSYNAHKKQKRSIADVRNEEQNIQNHTEL 230
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 171 -----EDEDDADLSKYN-----LDASEEDSNKKKSNRRS----- 204
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 231 QEKSSNATSKESPAPLHRRHMRHRLHVRKARSDSRPAAHFLSSRRRHQS 287
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 205 KSRSHSRSSRSPSSSRSRSRSRSSSSSSSSSSSSSSSS-----RERSRSGS 254
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

Search completed: September 13, 2004, 10:34:38





GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 13, 2004, 10:25:38 ; Search time 43 Seconds  
(without alignments)  
3001.092 Million cell updates/sec

Title: US-09-813-329-6

Perfect score: 2162

Sequence: 1 MTATLKPFTTTSANDGF.....DRNAVLREGNRRSYGIFKV 409

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL\_25:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organalle:\*
- 9: sp\_phase:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_virus:\*
- 16: sp\_bacteriap:\*
- 17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2162	100.0	409	Q8MY88	Q8MY88 drosophila
2	2157	99.8	409	Q8IGD3	Q8IGD3 drosophila
3	2149	99.4	415	Q8MUJ1	Q8MUJ1 drosophila
4	1656	76.6	325	Q9V5G2	Q9V5G2 drosophila
5	1359	62.9	261	Q8MRW2	Q8MRW2 drosophila
6	175.5	8.1	2226	Q97225	Q97225 plasmodium
7	172.5	8.0	2221	Q81259	Q81259 plasmodium
8	168	7.8	10061	Q81321	Q81321 plasmodium
9	165.5	7.7	3268	Q30280	Q30280 saccharomyc
10	164	7.6	2026	Q81565	Q81565 plasmodium
11	160.5	7.4	2309	Q81517	Q81517 plasmodium
12	159	7.4	491	Q81J07	Q81J07 plasmodium
13	154	7.1	2738	Q81HU4	Q81HU4 plasmodium
14	153.5	7.1	1371	Q81CK4	Q81CK4 plasmodium
15	153.5	7.1	3401	Q81BQ3	Q81BQ3 plasmodium
16	151	7.0	3452	Q81EA3	Q81EA3 plasmodium

#### ALIGNMENTS

RESULT 1

Q8MY88 ID Q8MY88 PRELIMINARY; PRT; 409 AA.  
AC Q8MY88;  
DT 01-OCT-2002 (TReMBLrel. 22, Created)  
DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)  
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)  
DE TNF superfamily ligand, Eiger (Tumor necrosis factor family member DTI).  
GN EIGER OR DT1 OR CG12919.  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=22060500; PubMed=12065414;  
RA Igaki T., Kanda H., Yamamoto-Goto Y., Kanuka H., Kuranaga E.,  
RA Aigaki T., Miura M.;  
RT "Eiger, a TNF superfamily ligand that triggers the Drosophila JNK pathway.";  
RT RT  
RL EMBO J. 21:3009-3018(2002).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Inohara N., Nunez G.;  
RT "DTI, a Drosophila tumor necrosis factor family member.";  
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AB073865; BAC00950.1; -  
DR EMBL; AF149799; AA015310.1; -  
DR Flybase; FBgn0033483; eiger.  
DR GO; GO:0016020; C:membrane; IEA.  
DR GO; GO:0005164; F:tumor necrosis factor receptor binding; IEA.  
DR GO; GO:0006955; P:immune response; IEA.  
DR InterPro; IPR006052; TNF family.  
DR InterPro; IPR008983; TNF\_like.  
DR SMART; SM00207; TNF; 1.  
DR PROSITE; PS00251; TNF\_1; 1.  
DR PROSITE; PS50049; TNF\_2; 1.

Q8IM62 plasmodium  
Q8IE75 plasmodium  
Q9QZV4 mus musculus  
Q8I2R3 plasmodium  
Q8I2D2 plasmodium  
Q9WVE4 mus musculus  
Q8I0W7 plasmodium  
Q8IJQ2 plasmodium  
Q8I4Y8 plasmodium  
Q12500 saccharomyc  
Q8I298 plasmodium  
Q8IDY3 plasmodium  
Q8I1D5 plasmodium  
Q8I0E6 schistosoma  
Q9U0L0 plasmodium  
Q8I350 plasmodium  
Q8I1B1 plasmodium  
Q8I1R1 plasmodium  
Q8MB1 dictyosteli  
Q8IM32 plasmodium  
Q8IKX1 plasmodium  
Q8I3R5 plasmodium  
Q86127 plasmodium  
Q8ILU2 plasmodium  
Q24376 drosophila  
Q61143 plasmodium  
Q8IE99 plasmodium  
Q8IKF6 plasmodium  
Q8I2K4 plasmodium



Query Match	76.6%	Score 1656;	DB 5;	Length 325;
Best Local Similarity	97.8%	Pred. No. 7.1e-103;		
Matches 310;	Conservative 1;	Mismatches 0;	Indels 6;	Gaps 1
QY	99	EYENALIDYPKKVDGLTDEBDDDDGDLGDSIADDEDDDDVSYSVDDVGADYEDYTDMLNK	158	
Db	9	QYENALIDYPKKVDGLTDEBDDDDGDLGDSIADDEDDDDVSYSVDDVGADYEDYTDMLNK	68	
QY	159	LNNAHGTGTPPTSETTASGEGETDSASASNDNVDFDFTSYNAHKKKQKRSRSIADVRN	218	
Db	69	LNNAHGTGTPPTSETTASGEGETDSASASNDNVDFDFTSYNAHKKKQKRSRSIADVRN	128	
QY	219	EEQNIQGNHTELOEKSNEATSXESAPAPLHRRRMRHSRHLHLVRK-----ARSEDSP	272	
Db	129	EEQNIQGNHTELOEKSNEATSXESAPAPLHRRRMRHSRHLHLVRKESILLSARSEDSP	188	
QY	273	AAHFHLSRRRHQSGSMGYHGMVIGNDNERNYSQGHFQTRDGVLTVTNTGLYYVVAQICY	332	
Db	189	AAHFHLSRRRHQSGSMGYHGMVIGNDNERNYSQGHFQTRDGVLTVTNTGLYYVVAQICY	248	
QY	333	NNSHDQNGFVIFQGDTPFLOCLNTVPTNMPKHVHTCHTSLGLHLERNRIHLKDHNDRN	392	
Db	249	NNSHDQNGFVIFQGDTPFLOCLNTVPTNMPKHVHTCHTSLGLHLERNRIHLKDHNDRN	308	
QY	393	AVLREGNNRSYFGIFKV 409		
Db	309	AVLREGNNRSYFGIFKV 325		
RESULT 5				
Q8MRW2		PRELIMINARY;	PRT;	261 AA.
ID	Q8MRW2			
AC	Q8MRW2;			
DT	01-OCT-2002 (TremBLrel. 22, Created)			
DT	01-OCT-2002 (TremBLrel. 22, Last sequence update)			
DT	01-OCT-2003 (TremBLrel. 25, Last annotation update)			
DE	SD18286p.			
DE	EIGER OR CG12919.			
OS	Drosophila melanogaster (Fruit fly).			
OC	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;			
OC	Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;			
OC	Ephydroidea; Drosophilidae; Drosophila.			
OX	NCBI_TaxID=7227;			
RN	[1]			

```
RP SEQUENCE FROM N.A.
RA Stapleton M., Brokstein P., Hong L., Aghayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Drenken D., Farfan D., Frise E.,
RA George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,
RA Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Celniker S.,
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY119233; RAMS1093.1; -.
DR FlyBase; FBgn0033483; elgcr.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005164; F:tumor necrosis factor receptor binding; IEA.
DR GO; GO:0006955; P:immune response; IEA.
DR InterPro; IPR006052; TNF family.
DR InterPro; IPR008983; TNF-like.
DR SMART; SM00207; TNF; 1.
DR PROSITE; PS00251; TNF 1; 1.
DR PROSITE; PS0049; TNF 2; 1.
DR PROSITE; PS0049; TNF 2; 1.
SQ SEQUENCE 261 AA; 29780 MW; 13B6D5A04EC9122C CRC64;

Query Match      62.9%; Score 1359; DB 5; Length 261;
Best Local Similarity 97.7%; Pred.No. 3.8e-83;
Matches 255; Conservative 0; Mismatches 0; Indels 6; Gaps 1;

QY 155 MLNKLNNAAHTGTTPTSETTAGEGETDSASSASNDNDVDFDFTSYNAHKKKQERKRSIA 214
Db 1 MLNKLNNAAHTGTTPTSETTAGEGETDSASSASNDNDVDFDFTSYNAHKKKQERKRSIA 60

QY 215 DVNBEQNIQGNHTLQKSSNEATSKESAPLHRRMRHSRHLIVRK-----ARSE 268
Db 61 DVNBEQNIQGNHTLQKSSNEATSKESAPLHRRMRHSRHLIVRKGESLLSARSE 120

QY 269 DSRPAAHPLSSRRRHQSGMGVGHGDMYIGNDNERNYSQGHFQTRDGLVTWNTGLYYVYA 328
Db 121 DSRPAAHPLSSRRRHQSGMGVGHGDMYIGNDNERNYSQGHFQTRDGLVTWNTGLYYVYA 180

QY 329 QICYNNSHDQNGFIVFQGDTPFLQCLNTVPTNMPHKVHTCHTSLIHLERNERHLKDIH 388
Db 181 QICYNNSHDQNGFIVFQGDTPFLQCLNTVPTNMPHKVHTCHTSLIHLERNERHLKDIH 240

QY 389 NDRNAVLREGNRSYFGIFKV 409
Db 241 NDRNAVLREGNRSYFGIFKV 261

RESULT 6
O97225 ID O97225 PRELIMINARY; PRT; 2226 AA.
AC O97225;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein, conserved.
GN MAL3P2.2.
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=36329;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99376085; PubMed=10448855;
RA Bowman S., Lawson D., Basham D., Brown D., Chillingworth T.,
RA Churcher C.M., Craig A., Davies R.M., Devlin K., Feltwell T.,
RA Gentles S., Gilliam R., Hamlin N., Harris D., Holroyd S., Hornsby T.,
RA Horrocks P., Jagels K., Jassal B., Kyes S., McLean J., Moule S.,
RA Mungall K., Murphy J., Oliver K., Quail M.A., Rajandream M.-A.,
RA Rutter S., Skelton J., Squares R., Squares S., Sulston J.E.,
RA Whitehead S., Woodward J.R., Newbold C., Barrell B.G.;
RT "The complete nucleotide sequence of chromosome 3 of Plasmodium
falciparum .";
RL Nature 400:532-538 (1999).
[2]
RP SEQUENCE FROM N.A.
RX MEDLINE=22255708; PubMed=12368867;
```

```
RA Hall N., Pain A., Berriman M., Churcher C., Harris B., Harris D.,
RA Mungall K., Bowman S., Atkin R., Baker S., Barron A., Brooks K.,
RA Buckee C.O., Burrows C., Cherevach I., Chillingworth C.,
RA Chillingworth T., Christodoulou Z., Clark L., Clark R., Corton C.,
RA Cronin A., Davies R., Davis P., Dear P., Dearden F., Doggett J.,
RA Feltwell T., Goble A., Goodhead I., Gwilliam R., Hamlin N., Hance Z.,
RA Harper D., Hauser H., Hornsby T., Holroyd S., Horrocks P.,
RA Humphray S., Jagels K., James K.D., Johnson D., Kerhornou A.,
RA Knights A., Konfortov B., Kyes S., Larke N., Lawson D., Lennard N.,
RA Line A., Maddison M., McLean J., Mooney P., Moule S., Murphy L.,
RA Oliver K., Ormond D., Price C., Quail M.A., Rabinowitsch E.,
RA Rajandream M.A., Rutter S., Rutherford K.M., Sanders M., Simmonds M.,
RA Seeger K., Sharp S., Smith R., Squares R., Squares S., Stevens K.,
RA Taylor K., Tivey A., Unwin L., Whitehead S., Woodward J.,
RA Sulston J.E., Craig A., Newbold C., Barrell B.G.;
RT "Sequence of Plasmodium falciparum chromosomes 1, 3-9 and 13.";
RL Nature 419:527-531 (2002).
DR EMBL; AL034558; CAB38989.3; -.
DR InterPro; IPR007259; SPC97_Spc98.
DR Pfam; PF04130; SPC97_Spc98; 1.
KW Hypothetical protein.
SQ SEQUENCE 2226 AA; 267974 MW; 8690501ED4994768 CRC64;

Query Match      8.1%; Score 175.5; DB 5; Length 2226;
Best Local Similarity 21.9%; Pred.No. 0.0062;
Matches 82; Conservative 57; Mismatches 150; Indels 85; Gaps 16;

QY 69 KELSKLRVVDNLQORLGINYLDEFDEF-----QKEYENALIDYPKKVDGLTDEEDDDG 123
Db 155 KETKRKKRTLHNKN-----DTFGSVTSMNDEKDYMKYNQDYHDKDDDDYDEDDDED 207

QY 124 DGLDSTADDEDDVSYSSVDVGDYEDYDMLKLNNAHTGTTPTSETTAGEGETDSA 183
Db 208 D--DDYDDEDDDD--YEDDDDD--DDYDDEDDNNNSYDQNTGKKNKINPM 258

QY 184 SSASNDNDVDFDFTSYNAHKKKQERKRSIADVRN-----EQNIQGNHTLQK----- 232
Db 259 YNSTFETSNMFCVKDEKKEKQKNSI--DKRNIIYDSDSDNNYEHIFTHNSDLFF 317

QY 233 -KSSNEATSKESAPLHRRMRHSRHLIVRKARSDSRPAAHPLSSRRRHQSGMGVH 291
Db 318 LSIISNDHIEKEN--NLYIQNEQFINYDVFIRKFK-----FH 352

QY 292 GDMYIGNDNER-----NSYOGHFQTRDGLVTWNTGLYYV--YQAI--CYNNS 335
Db 353 KNLFVNMDSCNSWIVKKSILINETRSQKND--LVDINDVLYDVNMHRYNCYTRL 410

QY 336 HDQNGFIVFQGDTPFL-----QCLNTVPTNMPHKVHTCHTSLIHLERNERHLKDIH 388
Db 411 YINDDFLLNNKTKELGVQNVASHINNTNNNNMNNIDNIHNVNNIN--NYNREHLKRIE 469

QY 389 NDRNAVLREGNRS 402
Db 470 KKKKINNNYNNNN 483

RESULT 7
Q81259 ID Q81259 PRELIMINARY; PRT; 2221 AA.
AC Q81259;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein.
GN PFA0410W.
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=36329;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22255708; PubMed=12368867;
RA Hall N., Pain A., Berriman M., Churcher C., Harris B., Harris D.,
RA Mungall K., Bowman S., Atkin R., Baker S., Barron A., Brooks K.,
```





RESULT 11		
Q81517	PRELIMINARY;	PRT; 2309 AA.
ID		
AC	Q81517	
DT	01-MAR-2003 (TrEMBLrel. 23, Created)	
DT	01-MAR-2003 (TrEMBLrel. 23, Last sequence update)	
DT	01-MAR-2003 (TrEMBLrel. 23, Last annotation update)	
DE	Hypothetical protein.	
GN	PFL1980C.	
OS	Plasmodium falciparum (isolate 3D7).	
OC	Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.	
OX	NCBI_TaxID=36329;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=3D7;	
RX	MEDLINE=22255705; PubMed=12368864;	
RA	Gardner M.J., Hall N., Fung E., White O., Berriman M., Hyman R.W.,	
RA	Carlton J.M., Paten A., Nelson K.E., Bowman S., Paulsen I.T., James K.,	
RA	Eisen J.A., Rutherford K., Salzberg S.L., Craig A., Kyes S.,	
RA	Chan M.-S., Nene V., Shallow S.J., Suh B., Peterson J., Angiuoli S.,	
RA	Pertea M., Allen J., Salengut J., Haft D., Mather M.W., Vaidya A.B.,	
RA	Martin D.M.A., Fairlamb A.H., Fraunholz M.J., Roos D.S., Ralph S.A.,	
RA	McFadden G.I., Cummings L.M., Subramanian G.M., Mungall C.,	
RA	Venter J.C., Carucci D.J., Hoffman S.L., Newbold C., Davis R.W.,	
RA	Fraser C.M., Barrell B.;	
RT	"Genome sequence of the human malaria parasite Plasmodium	
RT	falciparum.";	
RL	Nature 419:498-511(2002).	
DR	EMBL; AE014850; AAN36480.1; --	
DR	Hypothetical protein.	
SK	SEQUENCE 2309 AA; 276528 MW; 8A6F8DBB972C4922 CRC64;	
<p>Query Match 7.4%; Score 160.5; DB 5; Length 2309;            Best Local Similarity 19.6%; Pred. No. 0.066;            Matches 70; Conservative 77; Mismatches 118; Indels 93; Gaps 15;</p>		
Qy	69 KELKSLKRVVDNLQORLGINLYLDFD--EFOKEYENALIDYPKVDGLTDEEDDDGDL 126	
Db	701 KYLMKYKNVSNLQ-----DFNKDKYSESEGEVNDPVKL-----DQSNINSRGS 747	
Qy	127 DSIADDEDDVYSVSDVGA-DYEDYTD-----MLKLNNAHTGTPPTSETTAEGE 179	
Db	748 YTSHYNKGVDNTSSYDKNSYSDDDYDTSKFQENRYKHSYK-NVYKNKSSDDYEND 806	
Qy	180 TDSASSASNDNVF--DDFTSYNAHKQKQERKSIADVRNEQNIQGNHTLEOKS--- 234	
Db	807 SNSSNNNNNNNTYSDDDFNSYNNRRNTQKSKQMKNERQKNKKKKNKIDQSDSTWD 866	
Qy	235 -----SNEATSKESPAPLHRRHRMHSRHRHLVKARSEDSPAAHPFLHSRRHQ 285	
Db	867 EYNNMNDLYSNEYNRSIFNDILHAKKRFYDLISINRMKNT----- 911	
Qy	286 GSMGYHGDYMGNDNERNYSOGHFQTRDGVLTVTNTGLYYVYAQICYNNSHDQNGFIVQ 345	
Db	912 -----FNVNNINKNKTID--FNFIENIDTIS---YFT---SIYNSEH-----IFQ 948	
Qy	346 GDTPFLOCLNTVPTNMPKHVHTCHTSGLIHLERNERHLKDIHNDRAVLREGNRSY 403	
Db	949 FDYFLFAIEICTYNNYI-----KSYLINNNNN---NNNNLKY 988	
<p>RESULT 12</p>		
Q81JQ7	PRELIMINARY;	PRT; 491 AA.
ID	Q81JQ7	
AC	Q81JQ7	
DT	01-MAR-2003 (TrEMBLrel. 23, Created)	
DT	01-MAR-2003 (TrEMBLrel. 23, Last sequence update)	
DT	01-MAR-2003 (TrEMBLrel. 23, Last annotation update)	
DE	Hypothetical protein.	
GN	Pf10_0135.	
OS	Plasmodium falciparum (isolate 3D7).	
OC	Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.	











GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: September 15, 2004, 02:20:57 ; Search time 6712 Seconds  
(without alignments)  
2641.135 Million cell updates/sec

Title: US-09-813-329-6  
Perfect score: 2162  
Sequence: 1 MTATLTKPFIPTTSANDGF.....DRNAVLREGNNRSYGFIFKV 409

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 3470272 seqs, 21671516995 residues  
Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+p2n.model -DEV=xlp  
-Q=/cgn2\_1/USPTO\_spool\_p/US09813329/runat\_13092004\_102753\_21667/app\_query.fasta\_1.583  
-DB=genEmbl -QWMT=fastcap -SUFFIX=rgc -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US09813329 -CGEN 1 1 7406 @runat\_13092004\_102753\_21667 -NCPU=3  
-NO\_MMAP -LARGQUERY -NEG\_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl.\*  
1: gb.ba.\*  
2: gb.htg.\*  
3: gb.in.\*  
4: gb.om.\*  
5: gb.ov.\*  
6: gb.pat.\*  
7: gb.ph.\*  
8: gb.pl.\*  
9: gb.pr.\*  
10: gb.ro.\*  
11: gb.ats.\*  
12: gb.sy.\*  
13: gb.un.\*  
14: gb.vi.\*  
15: em.ba.\*  
16: em.fun.\*  
17: em.hum.\*  
18: em.in.\*  
19: em.mu.\*  
20: em.or.\*  
21: em.ov.\*  
22: em.pat.\*  
23: em.ph.\*  
24: em.pl.\*  
25: em.ro.\*  
26: em.ats.\*  
27: em.ats.\*  
28: em.un.\*

29: em.vi.\*  
30: em.htg.hum.\*  
31: em.htg.inv.\*  
32: em.htg.other.\*  
33: em.htg.mus.\*  
34: em.htg.pln.\*  
35: em.htg.rod.\*  
36: em.htg.mam.\*  
37: em.htg.vrt.\*  
38: em.gy.\*  
39: em.htgo.hum.\*  
40: em.htgo.mus.\*  
41: em.htgo.other.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	2162	100.0	1656	3	AB073865	Drosophil
2	2162	100.0	2101	3	AF149799	AF149799 Drosophil
3	2157	99.8	2165	3	BT001838	BT001838 Drosophil
4	2149	99.4	1248	3	AY115551	AY115551 Drosophil
5	2149	99.4	2159	3	AF521176	AF521176 Drosophil
6	1626	75.2	1221	3	AY119233	AY119233 Drosophil
7	1508.5	69.8	55359	2	AC012854	AC012854 Drosophil
c 8	1508.5	69.8	158983	3	AC005974	AC005974 Drosophil
c 9	1508.5	69.8	172904	3	AC007414	AC007414 Drosophil
c 10	1508.5	69.8	189620	3	AC099029	AC099029 Drosophil
c 11	1508.5	69.8	275390	3	AE003831	AE003831 Drosophil
12	175.5	8.1	158548	3	PFMAL3P2	AL034558 Plasmodiu
13	172.5	8.0	110000	3	PFMAL1P2_2	Continuation (3 of
14	168	7.8	313050	3	PFA929352	AL929352 Plasmodiu
15	165.5	7.7	69009	8	SCD8035	U33050 Saccharomyc
16	164	7.6	196490	2	AC005507	AC005507 Plasmodiu
c 17	164	7.6	250421	3	AE014849	AE014849 Plasmodiu
c 18	160.5	7.4	250713	3	AE014850	AE014850 Plasmodiu
c 19	160.5	7.4	256172	2	AC005139	AC005139 Plasmodiu
c 20	160.5	7.4	310779	2	AC005140	AC005140 Plasmodiu
c 21	154	7.1	110000	2	PFMAL13_08	Continuation (9 of
c 22	154	7.1	253151	3	AE014842	AE014842 Plasmodiu
c 23	154	7.1	318221	2	PFMAL13P3	AL049184 Plasmodiu
24	153.5	7.1	2810	4	RABHRCB	U05080 Rabbit hist
25	153.5	7.1	110000	2	PFMAL6P1_07	Continuation (8 of
26	153.5	7.1	110000	2	PFMAL7P1_07	Continuation (8 of
27	152	7.0	267206	2	AC131871	AC131871 Rattus no
c 28	152	7.0	269285	2	AC120096	AC120096 Rattus no
c 29	151.5	7.0	2270	10	AF158597	AF158597 Mus muscu
c 30	151.5	7.0	133661	14	U93872	U93872 Kaposi's sa
c 31	151.5	7.0	208632	2	AC073711	AC073711 Mus muscu
c 32	151	7.0	565	6	BD231127	BD231127 Shuffling
33	151	7.0	2616	6	AX642183	AX642183 Sequence
34	150.5	7.0	250029	3	AE014816	AE014816 Plasmodiu
35	149	6.9	5301	9	AF061189	AF061189 Homo sapi
c 36	148.5	6.9	110000	2	PFMAL13_09	Continuation (10 o
c 37	148.5	6.9	110000	2	PFMAL13_10	Continuation (11 o
c 38	148	6.8	250029	3	AE014839	AE014839 Plasmodiu
39	148	6.8	341050	3	PFA929357	AL929357 Plasmodiu
c 40	147.5	6.8	251762	3	AE014851	AE014851 Plasmodiu
c 41	147	6.8	5057	8	SCYLR114C	Z73286 S.cerevisia
c 42	147	6.8	15693	8	YSCJ9354	US3878 Saccharomyc
c 43	147	6.8	37639	8	SCCHX11LA	X89514 S.cerevisia
44	146	6.8	2407	10	AF132218	AF132218 Mus muscu
45	146	6.8	250195	3	AE014831	AE014831 Plasmodiu

ALIGNMENTS

RESULT 1



**LOCUS** AF149799 2101 bp mRNA linear INV 02-JAN-2003  
**DEFINITION** Drosophila melanogaster tumor necrosis factor family member DT1  
**ACCESSION** AF149799  
**VERSION** AF149799.1 GI:27462085  
**KEYWORDS** Drosophila melanogaster (fruit fly)  
**SOURCE** Drosophila melanogaster  
**ORGANISM** Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.  
**REFERENCE** 1 (bases 1 to 2101)  
**AUTHORS** Inohara, N. and Nunez, G.  
**TITLE** DT1, a Drosophila tumor necrosis factor family member  
**JOURNAL** Unpublished  
**REFERENCE** 2 (bases 1 to 2101)  
**AUTHORS** Inohara, N. and Nunez, G.  
**TITLE** Direct Submission  
**JOURNAL** Submitted (10-MAY-1999) Department of Pathology, Comprehensive Cancer Center, 1500 E. Medical Center Dr., Ann Arbor, MI 48109, USA  
**FEATURES** Location/Qualifiers  
 1..2101  
     /organism="Drosophila melanogaster"  
     /mol\_type="mRNA"  
     /db\_xref="taxon:7227"  
     1..2101  
     /gene="dt1"  
     625..1854  
     /gene="dt1"  
     /codon\_start=1  
     /product="tumor necrosis factor family member DT1"  
     /protein\_id="AAO15310.1"  
     /db\_xref="GI:27462086"  
     /translation="MTAEELKPFITPTSSANDDGPFAKATSTATAQRETROLIPVLGF  
     IGLGIVAILALTWTQTRVSHLDKELSKLVVDNLQORLGINLYDEFDFQKEYEN  
     ALIDYPPKVDLGGDDDDGLDSDADDDDDVSYSSVDVGDVYEDYDMLKLN  
     NAHTGTTPTSETTAEGEGEDTSASASNDNDVDPDFTSYNAHKKKQKRSIADVRN  
     EEQNIQGNHTLQKSNSEATSKESPAPLHRRMRHSHRHLVYRKARSDSRPAAHF  
     HLSRRHQSMGVKDMYIGINDERNYSQGHFOTRDGVLTVTNGLVYVYVAQICYN  
     SHDQNGFIVFGDTPFLQCLNTVPTNNPKHVHTCHTSLIHLERNERILKDIHNDNRN  
     AVLREGNRSYFGIFKV"  
     2080..2085  
     /gene="dt1"  
     2101  
     /gene="dt1"  
**polYA\_signal**  
**polYA\_site**  
**ORIGIN**  
 Alignment Scores:  
 Pred. No.: 2,65e-142 Length: 2101  
 Score: 2162.00 Matches: 409  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 3 Gaps: 0  
 US-09-813-329-6 (1-409) x AF149799 (1-2101)  
 Qy 1 MetThrAlaGluThrLeuLysProPheIleThrProThrSerAlaAsnAspGlyPhe 20  
 Db 625 ATGACTGCCGAGACCCCTCAACCGCTTTATACGCCAACGAGTGCCACGATGATGGTTT 684  
 Qy 21 ProAlaLysAlaThrSerThrAlaThrAlaGlnArgArgThrArgGlnLeuIleProLeu 40  
 Db 685 CCGGCCAAGGACGACGACGCGACGCCCGCAGCGACCGCCGCGCTGATCCCCCTG 744  
 Qy 41 ValLeuGlyPheIleGlyLeuGlyLeuValValAlaIleLeuAlaLeuThrIleTrpGln 60  
 Db 745 GTTTTGGGCTTCATCGGCTCGGGCTGGTCTGCTTGCATTCTCGCACTAACCATCTGGCAG 804  
 Qy 61 ThrThrArgValSerHisLeuAspLysGluLeuLysSerLeuLysArgValValAspAsn 80  
 Db 805 ACAACGCGGTGATCGCATCTCGCAAGGAGCTGAAGAGCCTGAAGCGAGTCTGCGATAAT 864

Qy 81 LeuGlnGlnArgLeuGlyIleAsnTyrLeuAspGluPheAspGluPheGlnLysGluTyr 100  
 Db 865 CTCACGACGCGTTTGGGCATAAATACTCTGACGAGTTCGACGAGTTCCAAAAGAGGTAC 924  
 Qy 101 GluAsnAlaLeuIleAspTyrProLysValAspGlyLeuThrAspGluGluAspAsp 120  
 Db 925 GAGAATGCCCTCATCGACTATCCAAAAAAGGTGGATGGCTCTACGATGAGGAGACGAC 984  
 Qy 121 AspAspGlyAspGlyLeuAspSerIleAlaAspAspGluAspAspValSerTyrSer 140  
 Db 985 GACATGCGATGCTCTGGATTCCATTGCGACGACGAGACGACGACGCTTAGCTATAGC 1044  
 Qy 141 SerValAspAspValGlyAlaAspTyrGluAspTyrThrAspMetLeuAsnLysLeuAsn 160  
 Db 1045 TCTGTGATGATGTTGGCGCAGACTACGAGGACTACACCGATATGTTAAATAACTCAAC 1104  
 Qy 161 AsnAlaHisThrGlyThrThrProThrSerGluThrThrAlaGluGlyGluGlyGluThr 180  
 Db 1105 AATGCACATACCGGCACACCGCCACATCTGAGACCACTGCTGAGGCGGAGGCGAGACG 1164  
 Qy 181 AspSerAlaSerSerAlaSerAsnAspAspValPheAspAspPheThrSerTyrAsn 200  
 Db 1165 GACATGCACTCTCAGCGCTCAATATGATGACATGTGTTCGATGACTTTTACGAGCTACAT 1224  
 Qy 201 AlaHisLysLysLysGlnGluArgLysSerArgSerIleAlaAspValArgAsnGluGlu 220  
 Db 1225 GCCCACAAAAAAGACGAGGAGAGAAAATCTCGCTCGATTGCCGATGTACGCAATGAGGAG 1284  
 Qy 221 GlnAsnIleGlnGlyAsnHisThrGluLeuGlnGlyLysSerSerAsnGluAlaThrSer 240  
 Db 1285 CAGATAATTCAAGGAAATCAACAGAGCTTCAGGAAAGTCAATCCATAGAGGCAACTTCC 1344  
 Qy 241 LysGluSerProAlaProLeuHisArgArgArgMetHisSerArgHisArgHisLeu 260  
 Db 1345 AAAGAGAGCCCTGCACCACTTACCACCGTCGAGATGATGATTTCCGCCCATCGCCACTC 1404  
 Qy 261 LeuValArgLysAlaArgSerGluAspSerArgProAlaAlaHisPheHisLeuSerSer 280  
 Db 1405 CTAGTCGCAAAAGCCAGATCCGAGGACTCGAGGCGCAGCCCACTTTTCCACTTCGAGCAGC 1464  
 Qy 281 ArgArgArgHisGlnGlySerMetGlyTyrHisGlyAspMetTyrIleGlyAsnAspAsn 300  
 Db 1465 AGCGCGGTCAACCAAGGAAGTATGGGCTACCATGGAGATATGTACATAGGAAATGATAAC 1524  
 Qy 301 GluArgAsnSerTyrGlnGlyHisPheGlnThrArgAspGlyValLeuThrValThrAsn 320  
 Db 1525 GAGAAAAACTTTATCAGGACACTTTCAAACCGCGCATGGCGTCTTGACGCTGACCAAT 1584  
 Qy 321 ThrGlyLeuTyrTyrValTyrAlaGlnIleCysTyrAsnAsnSerHisAspGlnAsnGly 340  
 Db 1585 ACAGGCTATATTACGTATACGCCCGCAGATATGCTACAACAACTCGCAGCAGCAGACGGA 1644  
 Qy 341 PheIleValPheGlnGlyAspThrProPheLeuGlnCysLeuAsnThrValProThrAsn 360  
 Db 1645 TTTATCGTCTTCAAAGGAGACACTCCATTCTCGAGTGTGTAACACGCGTGCCCAACCAAC 1704  
 Qy 361 MetProHisLysValHisThrCysHisThrSerGlyLeuIleHisLeuGluArgAsnGlu 380  
 Db 1705 ATGCCAATAAAGGTGCACACTGCCACACGAGTGGTCTGATCCACTTGGAAACGAAACGAG 1764  
 Qy 381 ArgIleHisLeuLysAspIleHisAsnAspArgAsnAlaValLeuArgGluGlyAsnAsn 400  
 Db 1765 AGGATCCATCTGAAGGACATTCACACGATCGCAATGCACTTCTTCGCGGGAGGAAACAAC 1824  
 Qy 401 ArgSerTyrPheGlyIlePheLysVal 409  
 Db 1825 CGAAGCTACTTTGGCATCTTCAAGGTG 1851  
 RESULT 3  
 BT001838  
 LOCUS BT001838 2165 bp mRNA linear INV 15-NOV-2002  
 DEFINITION Drosophila melanogaster RH51659 full insert cDNA.  
 ACCESSION BT001838

VERSION	BT001838.1	GI:25013017	
KEYWORDS	FLI CDNA		
SOURCE	Drosophila melanogaster (fruit fly)		
ORGANISM	Drosophila melanogaster		
REFERENCE	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.		
AUTHORS	1 (bases 1 to 2165) Stapleton, M., Brokstein, P., Hong, L., Agbayani, A., Carlson, J., Champe, M., Chavez, C., Dorsett, V., Dresnek, D., Farfan, D., Frise, E., George, R., Gonzalez, M., Guarin, H., Krommiller, B., Li, P., Liao, G., Miranda, A., Mungall, C. J., Nunoo, J., Pacleb, J., Paragae, V., Park, S., Patel, S., Phouanavong, S., Wan, K., Yu, C., Lewis, S. E., Rubin, G. M. and Celnikier, S.		
TITLE	Direct Submission		
JOURNAL	Submitted (15-NOV-2002) Berkeley Drosophila Genome Project, Lawrence Berkeley National Laboratory, One Cyclotron Road, Berkeley, CA 94720, USA		
COMMENT	Sequence submitted by: Berkeley Drosophila Genome Project Lawrence Berkeley National Laboratory Berkeley, CA 94720 This clone was sequenced as part of a high-throughput process to sequence clones from Drosophila Gene Collection 1 (Rubin et al., Science 2000). The sequence has been subjected to integrity checks for sequence accuracy, presence of a polyA tail and contiguity within 100 kb in the genome. Thus we believe the sequence to reflect accurately this particular cDNA clone. However, there are artifacts associated with the generation of cDNA clones that may have not been detected in our initial analyses such as internal priming, priming from contaminating genomic DNA, retained introns due to reverse transcription of unspliced precursor RNAs, and reverse transcriptase errors that result in single base changes. For further information about this sequence, including its location and relationship to other sequences, please visit our Web site ( <a href="http://fruitfly.berkeley.edu">http://fruitfly.berkeley.edu</a> ) or send email to <a href="mailto:cdna@fruitfly.berkeley.edu">cdna@fruitfly.berkeley.edu</a> .		
FEATURES	Location/Qualifiers		
source	1..2165 /organism="Drosophila melanogaster" /mol_type="mRNA" /strain="y; cn bw sp" /db_xref="taxon:7227"		
misc_feature	1..2152 /note="sim4 alignment with AE003831 (46B13-46B4)"		
CDS	648..1877 /note="Longest ORF" /codon_start=1 /product="RHS1659p" /protein_id="AA071595.1" /db_xref="GI:25013018" /translation="MTAETLKPFTTPTSSANDGFPKATSTATAQRRTQLPLVLGF IGLGVAILALITWOTRYSHLDKELSLKRVVDNLQRLGINLYLDEDEQKEYEN ALIDYKPKVGLTDEDDDDGGLDSIADDEDDVSVDVGDYEDYTDMLNKL NAHTGPTTSETTAEGGETDSASSNDUNVDFDFTSNNAHKKQKERSKIADVRN EEQIQNHTLEQKSSGMSKATSKERPAFLHRRMHSRHLVVRKARSEDGRPAHF HLSRRHQSGMGHDYITGNDNRNSYQGFQTRDGLVTNTGLYVYVAQICVNN SHDQNGFIVFGDTPFLQCLNTVPTNMPKHVTCHTSLGLIHLERNERIHLDIHDRN AVLRGENNRSYFGIFKY"		
ORIGIN			
Alignment Scores:			
Pred. No.:	6.13e-142	Length:	2165
Score:	2157.00	Matches:	408
Percent Similarity:	99.76%	Conservative:	0
Best Local Similarity:	99.76%	Mismatches:	1
Query Match:	99.77%	Indels:	0
DB:	3	Gaps:	0
US-09-813-329-6 (1-409) x BF001838 (1-2165)			
Qy	1 MetThrAlaGluThrLeuLysProPheIleThrProThrSerAlaAsnAspGlyPhe		20
	648 ATGACTCCCGAGACCCCTCAAGCCGTTTATAACCCCAACGAGTGCCCAACGATGATGTTTT		707
Qy	21 ProAlaLysAlaThrSerThrAlaThrAlaGlnArgArgThrArgGlnLeuIleProLeu		40
Db	708 CCGGCCAAAGGACCGACCGCGACCGCCACGAGCGACCCGCCAGTGATGCCCTG		767
Qy	41 ValLeuGlyPheIleGlyLeuGlyLeuValValAlaIleLeuAlaLeuThrIleTrpGln		60
Db	768 GTTTTGGGGTTCATCGGTCTGGGGCTGGTCTGTTCGCACTTCTCGCACTAACGATCGCAG		827
Qy	61 ThrThrArgValSerHisLeuAspLysGluLeuLysSerLeuLysArgValValAspAsn		80
Db	828 ACAACCGGTGATCGCATCTGGACAAGGAGCTGAAGAGCTGAAGCGAGTCTCGTCAATA		887
Qy	81 LeuGlnGlnArgLeuGlyIleAsnTyrLeuAspGluPheAspGluPheGlnLysGluTyr		100
Db	888 CTCACGACGCGTTTGGGCATAAACTATCTGGACGAGTTCGACGAGTTCANAAAGGAGTAC		947
Qy	101 GluAsnAlaLeuIleAspTyrProLysLysValAspGlyLeuThrAspGluAspAsp		120
Db	948 GAGATGCCCTCATTCGACTATCCAAAAAGGTGGATGGCTCACGATGAGGAGACGAC		1007
Qy	121 AspAspGlyAspGlyLeuAspSerIleAlaAspAspGluAspAspValSerTyrSer		140
Db	1008 GACGATGGCGATGGTCTGGATTCCATTGCGGACGACGAGACGACGCTTAGCTATAGC		1067
Qy	141 SerValAspAspValGlyAlaAspTyrGluAspTyrThrAspMetLeuAsnLysLeuAsn		160
Db	1068 TCTGTGGATGATGTTGGCGCAGACTACGAGGACTACCCGATATGTAAATAAATCACTCA		1127
Qy	161 AsnAlaHisThrGlyThrThrProThrSerGluThrThrAlaGluGlyGluGlyThr		180
Db	1128 AATGCATACACCGGACCAACGCCCATCTGAGACCACTGCTGAGGGCGGAGGCGAGAGC		1187
Qy	181 AspSerAlaSerSerAlaSerAsnAspAspValPheAspAspPheThrSerTyrAsn		200
Db	1188 GACAGTGCATCTCAGCCTCAATGATGACATGTTGTCGATGACTTTTACGAGCTACAA		1247
Qy	201 AlaHisLysLysLysGlnGluArgLysSerArgSerIleAlaAspValArgAsnGlu		220
Db	1248 GCCCAAAAGAGAACGACGAGAGAAAATCTCGTCGATGCGCATGTACGCAATGAGGAG		1307
Qy	221 GlnAsnIleGlnGlyAsnHisThrGluLeuGlnGluLysSerSerAsnGluAlaThrSer		240
Db	1308 CAGATATTCAAGGAATACACAGAGCTTCAGGAAAAGTATCATCAATGAGGCACTCC		1367
Qy	241 LysGluSerProAlaProLeuHisHisArgArgArgMetHisSerArgHisArgHisLeu		260
Db	1368 AAAGAGAGACCTGCACCACTTCAACCCGTCGAGATGTCATTCGCGCATCGCACCTC		1427
Qy	261 LeuValArgLysAlaArgSerGluAspSerArgProAlaAlaHisPheHisLeuSerSer		280
Db	1428 CTAGTCCGCAAGGACAGATCCGAGACTCGAGGCGCAGAGCCCATTTCCACTTGAGCAGC		1487
Qy	281 ArgArgArgHisGlnGlySerMetGlyTyrHisGlyAspMetTyrIleGlyAsnAspAsn		300
Db	1488 AGGCGCGTCAACCAAGGAATATGGCTACCATGGAGATATGATACATAGGAAATGATAAC		1547
Qy	301 GluArgAsnSerTyrGlnGlyHisPheGlnThrArgAspGlyValLeuThrValThrAsn		320
Db	1548 GAGAGAACTCTTATCAGGGACACTTTCAACCGCGATGGCGTCTTGACGCTGACCAAT		1607
Qy	321 ThrGlyLeuTyrTyrValTyrAlaGlnIleCysTyrAsnAsnSerHisAspGlnAsnGly		340
Db	1608 ACAGGCTTATATTAGTATACGCCAGATATGCTACAACTCGCAGCACCAGAACCGGA		1667
Qy	341 PheIleValPheGlnGlyAspThrProPheLeuGlnCysLeuAsnThrValProThrAsn		360
Db	1668 TTTATCGTCTTTCAAGGAGACACTCCCATTCCTGCGAGTGTGTAACACGCTGCCACCAAC		1727
Qy	361 MetProHisLysValHisThrCysHisThrSerGlyLeuIleHisLeuGluArgAsnGlu		380
Db	1728 ATGCCATAGGTGCACACCTGTCACACGAGTGTCTGTATCCACCTGGAAACGAAACGAG		1787

```

QY 381 ArgIleHisLeuLysAspIleHisAenAspArgAenAlaValLeuArgGluGlyAsnAen 400
|
|
|
Db 1788 AGGATCCATCTGAAGGACATTACAAACGATCGCATGCGAGTCTTCGCGGAGGGAACAC 1847
|
|
|
QY 401 ArgSerTyrPheGlyIlePheLysVal 409
|
|
|
Db 1848 CGAAGCTACTTTGGCATCTTCAAGGTG 1874

RESULT 4
AY115551
LOCUS AY115551 1248 bp mRNA linear INV 01-AUG-2003
DEFINITION Drosophila melanogaster DARTH (darth) mRNA, complete cds.
ACCESSION AY115551
VERSION AY115551.1 GI:31321973
KEYWORDS
SOURCE
ORGANISM
Drosophila melanogaster (fruit fly)
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 1248)
Chauppila,S., Maaty,W.S., Chen,P., Tomar,R.S., Eby,M.T., Chapo,J.,
Chew,S., Rathore,N., Zachariah,S., Sinha,S.K., Abrams,J.M. and
Chaudhary,P.M.
Eiger and its receptor, Wengen, comprise a TNF-like system in
Drosophila
Oncogene 22 (31), 4860-4867 (2003)
JOURNAL
MEDLINE 22759338
PUBMED 12894227
REFERENCE 2 (bases 1 to 1248)
AUTHORS Chaudhary,P.M.
TITLE Direct Submission
JOURNAL Submitted (28-MAY-2002) Internal Medicine, UT Southwestern Medical
Center, 5323 Harry Hines Blvd., Dallas, TX 75390, USA
FEATURES
source
1..1248
/organism="Drosophila melanogaster"
/mol_type="mRNA"
/db_xref="taxon:7227"
/chromosome="2"
/map="2R"
1..1248
/gene="darth"
1..1248
/gene="darth"
/note="Drosophila aspartate rich TNF homolog; tumor
necrosis factor ligand family member"
/codon_start=1
/product="DARTH"
/protein_id="AAM66763.1"
/db_xref="GI:31321974"
/translation="MTAEILKPFITPTTSANDGPPAKATSTATAQRTROLIPLVLGF
IGLGLVALIALTWTQTRVSHLDKELSKRVVDNLQQRILGNYLDFEFOKEYEN
ALIDYPKKVDLTDEDDGDGLDSIADDDDDSYSSVDDVGADYEDYDMLKLN
NAHGTPTSETTAEGEETDSASANDNDVDFDTSYNAHKKKQKRSIAADVRN
EEQNIQGNHTLEQKSNKATSKESPAPLHRRMRHSRHRLLVKGESLLSARSED
RPAAHFLHSRRSRQSGVHGDMYIGNDNERNYSQGHFQTRDGVLTWTGLYYVA
QICYNNSHDQGFIVQGDTPFLQCLNTVPTNMPHKVHTCHTSLHLERNERIHLDK
IHDRNAVLRGNRRSYFGIPKY"
1..108
/misc_feature
/gene="darth"
/note="Region: intracellular domain"
109..177
/misc_feature
/gene="darth"
/note="Region: transmembrane domain"
178..1245
/misc_feature
/gene="darth"
/note="Region: extracellular domain"
433..438
/misc_feature
/gene="darth"
/note="cleavage site"
ORIGIN

```

```

Alignment Scores:
Pred. No.: 1..19e-141 Length: 1248
Score: 2149.00 Matches: 409
Percent Similarity: 98.55% Conservative: 0
Best Local Similarity: 98.55% Mismatches: 0
Query Match: 99.40% Indels: 6
DB: 3 Gaps: 1

US-09-813-329-6 (1-409) x AY115551 (1-1248)

QY 1 MetThrAlaGluThrLeuLysProPheIleThrProThrSerAlaAenAspArgGlyPhe 20
|
|
|
Db 1 ATGACTCCGAGACCCCTCAAGCCGTTTATAACGCCAACGAGTGCACACGATGATGGTGT 60
|
|
|
QY 21 ProAlaLysAlaThrSerThrAlaThrAlaGlnArgThrArgGlnLeuLeuProLeu 40
|
|
|
Db 61 CCGGCCAAAGCGACCGACGCGCGACCGCCAGCGACGCCGCCAGCTGATCCCCCTG 120
|
|
|
QY 41 ValLeuGlyPheIleGlyLeuGlyLeuValValAlaIleLeuAlaLeuThrIleTrpGln 60
|
|
|
Db 121 GTTTTGGGGTTCATCGGTCTGGGGCTGGTTCGTTGCCATTCTCGCACTAACGATCTG 180
|
|
|
QY 61 ThrThrArgValSerHisLeuAspLysGluLeuLysSerLeuLysArgValValAspAen 80
|
|
|
Db 181 ACAACGGGTGTATCGCATCTGGACAAGGAGCTGAAGAGCCCTGAAGCGAGTCTCGATAAT 240
|
|
|
QY 81 LeuGlnGlnArgLeuGlyIleAenTyrLeuAspGluPheAspGluPheGlnLysGluTyr 100
|
|
|
Db 241 CTCACGACGGCTTTGGGCATAAACTATCTGCAGCAGTTCGACGAGTTCCTCAAAAGGAGTAC 300
|
|
|
QY 101 GluAenAlaLeuLeuAspTyrProLysLysValAspGlyLeuThrAspGluGluAspAen 120
|
|
|
Db 301 GAGNATCGCTCATCGATCTCCNAAAAGGTGGATGGCTCAGCGATGAGGAGGACGAC 360
|
|
|
QY 121 AspAspGlyAspGlyLeuAspSerIleAlaAspAspGluAspAspValSerTyrSer 140
|
|
|
Db 361 GACGATCGCATGCTCGGATTCATTTCGCGACGACGAGGACGACGACGCTAGCTATAGC 420
|
|
|
QY 141 SerValAspAspValGlyAlaAspTyrGluAspTyrThrAspMetLeuAenLysLeuAen 160
|
|
|
Db 421 TCTGTGGATGATGTTGGCGCAGACTACGAGGACTACACCGATATGTTAAATAAATAAC 480
|
|
|
QY 161 AsnAlaHisThrGlyThrThrProThrSerGluThrThrAlaGluGlyGluGlyThr 180
|
|
|
Db 481 AATGCACATACCGGCACACGCCACATCTCGACACCACTGCTGAGGGCGGGGCGAGACG 540
|
|
|
QY 181 AspSerAlaSerSerAlaSerAenAspAspValPheAspAspPheThrSerTyrAen 200
|
|
|
Db 541 GACAGTGCATCCTCAGCCTCAAATGATGACAAATGATGTTTCGATGACTTTTACCGCTACAAT 600
|
|
|
QY 201 AlaHisLysLysLysGlnLysSerArgSerIleAlaAspValArgAenGluGlu 220
|
|
|
Db 601 GCCCACAAAAGAAGCAGGAGAGAAAATCTCGCTCGATTGCCGATGATGACCAATGAGGAG 660
|
|
|
QY 221 GlnAenIleGlnGlyAenHisThrGluLeuGlnLysSerSerAenGluAlaThrSer 240
|
|
|
Db 661 CAGATATTTCAGGAATATCACAGAGCTTCAGGAAAAGTCATCCATGAGGCAACTTCC 720
|
|
|
QY 241 LysGluSerProAlaProLeuHisHisArgArgArgMetHisSerArgHisArgHisLeu 260
|
|
|
Db 721 AAAGAGAGCCCTGCACCACTTCACCACCGTCGAGAGATGATCTCCGCCCATCGCCACTC 780
|
|
|
QY 261 LeuValArgLys-----AlaArgSerGluAspSerArgProAlaAla 274
|
|
|
Db 781 CTAGTCCGCAAGGTGAATCTCTTCTTCAGCCAGATCCGAGGACTCGAGGCCAGCAGCC 840
|
|
|
QY 275 HisPheHisLeuSerSerArgArgHisGlnGlySerMetGlyTyrHisGlyAspMet 294
|
|
|
Db 841 CATTTCCATTTGAGCAGCAGCGCGGTCCACAGGAAGTATGGGCTACCATGAGATATG 900
|
|
|
QY 295 TyrIleGlyAenAspAenGluArgAenSerTyrGlnGlyHisPheGlnThrArgAspGly 314
|
|
|

```

```

Db      901  TACATAGGAATGATTAACGAGAGAACTCTTATCAGGACACTTTCAAACGGCGATGCC 960
Qy      315  ValLeuThrValThrAsnThrGlyLeuTyrTyrValTyrAlaGlnIleCysTyrAsnAsn 334
Db      961  GTCTTGACGGTACCAATAACAGCGCTATATTACGTATAGCTATAGCCAGATGCTTACAACAAC 1020
Qy      335  SerHisAspGlnAsnGlyPheIleValPheGlnGlyAspThrProPheLeuGlnCysLeu 354
Db      1021  TCGCAGCACAGACGAGATTTATCGTCTTTCAAGGAGACACTCCATTCTCGAGTGGCTTG 1080
Qy      355  AsnThrValProThrAsnMetProHisLysValHisThrCysHisThrSerGlyLeuIle 374
Db      1081  AACACGGTGGCCACCAACATGCCACATAAGGTGCACACTGCCACACAGTGGTCTGATC 1140
Qy      375  HisLeuGluArgAsnGluArgIleHisLeuLysAspIleHisAsnAspArgAsnAlaVal 394
Db      1141  CACCTGGGAACGAAACGAGAGGATCCATCTGAAGGACATTCAACAGATCGCAATCGAGTT 1200
Qy      395  LeuArgGluGlyAsnAsnArgSerTyrPheGlyIlePheLysVal 409
Db      1201  CTGCGGGAGGAAACAAACGGAAGTACTTTGGCATCTTCAAGGTG 1245

RESULT 5
AF521176      2159 bp      mRNA      linear      INV 14-AUG-2002
LOCUS      Drosophila melanogaster eiger mRNA, complete cds.
DEFINITION      AF521176
ACCESSION      AF521176
VERSION      AF521176.1  GI:21717645
KEYWORDS
SOURCE
ORGANISM      Drosophila melanogaster (fruit fly)
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 2159)
Moreno, E., Yan, M. and Basler, K.
Evolution of TNF Signaling Mechanisms. JNK-Dependent Apoptosis
Triggered by Eiger, the Drosophila Homolog of the TNF Superfamily
Curr. Biol. 12 (14), 1263-1268 (2002)
12176339
2 (bases 1 to 2159)
Moreno, E., Yan, M. and Basler, K.
Direct Submission
Submitted (13-JUN-2002) Institut Molecular Biology, Uni Zurich,
Winterthurerstrasse 190, Zurich, Z 8057, Switzerland
Location/Qualifiers
1. 2159
/organism="Drosophila melanogaster"
/mol_type="mRNA"
/db_xref="taxon:7227"
641..1888
/codon_start=1
/product="eiger"
/protein_id="AAW76710.1"
/db_xref="GI:21717646"
/translation="MTAETLKPFITPTSANDGFPKATSTATAQRRTQLIPLVLGF
IGLGVAILALATIWOTTRVSHLDKELSKRVVDNLQORLGINYLDEDFEPOKEVEN
ALIDYPKVDGLDEEDDDGDLSDIADDEDVSVSDVDVGADYEDYDMLNKLIN
NAHTGTTPTSETTAEGETDTSASSANDNDNVDDFTSYNAHKKQEKRSRIADVRN
EEQNIQGNHTLEOKSNEATSKESPAPLHRRMRSHRHLVVRKGSLLSARSEDS
RPAAHFLSKRRHQSGMYGHGDMYIGNDNERNYSOGHFQTRDGLVTNTGLYYVA
QICYNNSHDQNGFIVFGDTPFLQCLNTVPTNPHKVHTCHTSLGHLERNERIHLKD
IHNDRNAVLREGNRSYFGIFKY"

ORIGIN
Alignment Scores:
Pred. No.:      2,22e-141      Length:      2159
Score:          2149.00      Matches:      409
Percent Similarity: 98.55%      Conservative: 0
Best Local Similarity: 98.55%      Mismatches: 0
Query Match:     99.40%      Indels:      6
DB:              3          Gaps:      1

```

```

US-09-813-329-6 (1-409) x AF521176 (1-2159)
Qy      1  MetThrAlaGluThrIleLysProPheIleThrProThrSerAlaAsnAspAspGlyPhe 20
Db      641  ATGACTGCCGAGACCTTCAGCCGTTTATTAACCCAAACGAGTGCCACGATATGGTTTT 700
Qy      21  ProAlaLysAlaThrSerThrAlaThrAlaGlnArgThrArgGlnLeuIleProLeu 40
Db      701  CCGGCCAAAGCAGCAGCAGCGGCGACCGCCACGACGACCGCCAGCTGATCCCCCTG 760
Qy      41  ValIleGlyPheIleGlyLeuGlyLeuValValAlaIleLeuAlaLeuThrIleTyrGln 60
Db      761  GTTTTGGGGTTTCATCGGTCTGGGGCTGGTGGTTCATTCGCAATTCGCAATAACGATCTGGCAG 820
Qy      61  ThrThrArgValSerHisLeuAspLysGluLeuLysSerLeuLysArgValValAspAsn 80
Db      821  ACAACGGGTGTATCGCATCTGGACAGGAGCTGAGAGGCTGAAGCGATCTGTGATAAT 880
Qy      81  LeuGlnGlnArgLeuGlyIleAsnTyrLeuAspGluPheAspGluPheGlnLysGluTyr 100
Db      881  CTCACGACGCTTTGGGCATAAACTATCTGGACGAGTTCGACGAGTTCGAAAGGAGTAC 940
Qy      101  GluAsnAlaLeuIleAspTyrProLysLysValAspGlyLeuThrAspGluGluAspAsp 120
Db      941  GAGATGCCCTCATCGACTATCCAAAAAGGTGGATGGCTCACGGATGAGGAGACGAC 1000
Qy      121  AspAspGlyAspGlyLeuAspSerIleAlaAspAspGluAspAspValSerTyrSer 140
Db      1001  GACGATGGCGATGGTCTGGATTCCATTGGCAGCAGCAGGACGACGACGCTTAGCTATAGC 1060
Qy      141  SerValAspAspValGlyAlaAspTyrGluAspTyrThrAspMetLeuAsnLysLeuAsn 160
Db      1061  TCTGTGGATGATGTTGGCGCAGACTACGAGGACTACCGGATATGTTAAATAAATCAAC 1120
Qy      161  AsnAlaHisThrGlyThrThrProThrSerGluThrThrAlaGluGlyGluGlyGluThr 180
Db      1121  AATGCACATACCGGCACCGCCACATCTGAGACCATCTGTAGGGGCGAGGGCGAGAG 1180
Qy      181  AspSerAlaSerSerAlaSerAsnAspAsnValPheAspAspPheThrSerTyrAsn 200
Db      1181  GACAGTGCATCTCTAGCCTCAAATGATCAATGTGTTCGATGACTTTACCAGCTTACAAT 1240
Qy      201  AlaHisLysLysLysGlnGluArgLysSerArgSerIleAlaAspValArgAsnGluGlu 220
Db      1241  GCCCAAAAAGAACGACGAGAGAAAATCTCGCTCGATTGCGGATGTACGCAATGAGAGAG 1300
Qy      221  GlnAsnIleGlnGlyAsnHisThrGluLeuGlnLysSerSerAsnGluAlaThrSer 240
Db      1301  CAGAATATTCAAGGAAATCACACAGAGCTTCAGGAAAAGTCAATCCAATGAGGCAACTTCC 1360
Qy      241  LysGluSerProAlaProLeuHisHisArgArgArgMetHisSerArgHisArgHisLeu 260
Db      1361  AAAGAGAGCCCTGCACCACTTCAACCACGTCGCAAGATGCAATTCGCGCATCGCCACCTC 1420
Qy      261  LeuValArgLys-----AlaArgSerGluAspSerArgProAlaAla 274
Db      1421  CTAGTCCCAAGGTGAATCTCTCTTTCAGCCAGATCCGAGGACTCGAGGCCAGACGCC 1480
Qy      275  HisPheHisLeuSerSerArgArgHisGlnGlySerMetGlyTyrHisGlyAspMet 294
Db      1481  CATTTCCACCTCAGCAGCAGCGCGCGTCCACCACCAAGAGTATGGGTACCATGGAGATATG 1540
Qy      295  TyrIleGlyAsnAspAsnGluArgAsnSerTyrGlnGlyHisPheGlnThrArgAspGly 314
Db      1541  TACATAGGAATGATTAACGAGAGAACTCTTATCAGGAGACACTTTCAAACCGCGGATGGC 1600
Qy      315  ValLeuThrValThrAsnThrGlyLeuTyrTyrValTyrAlaGlnIleCysTyrAsnAsn 334
Db      1601  GTCTTGACGGTGACCAATACAGGCTATATTACGTATAGCCCCAGATATGCTACAAACAAC 1660
Qy      335  SerHisAspGlnAsnGlyPheIleValPheGlnGlyAspThrProPheLeuGlnCysLeu 354
Db      1661  TCGCACGACCAAGAACGGATTTATCGTCTTTCAAGGAGACACTCCATTTCTCTGCGAGTGTG 1720

```



QY 355 AnThrValProThrAsnMetProHisLysValHisThrCysHisThrSerGlyLeuile 374  
 Db 1721 AACCGGTGCCCAACATGACATGAAGGTGCACACCTGCCACACGAGTGTCTGATC 1780

QY 375 HisLeuGluArgAsnGluArgIleHisLeuLysAspIleHisAsnAspArgAsnAlaVal 394  
 Db 1781 CACCTGGAAACAAACAGAGGATCCATCTGAAGGACATTCACACGATCGCAATGCAGTT 1840

QY 395 LeuArgGluGlyAsnAsnArgSerTyrPheGlyIlePheLysVal 409  
 Db 1841 CTGGGGAGGGAACACCGAAGCTACTTTGGCATCTTCAAGGTG 1885

## RESULT 6

AY119233 1221 bp mRNA linear INV 16-JUN-2002  
 LOCUS Drosophila melanogaster SD18286 full insert cDNA.

DEFINITION AY119233  
 ACCESSION AY119233  
 VERSION AY119233.1 GI:21430829

KEYWORDS FLI CDNA.  
 SOURCE Drosophila melanogaster (fruit fly)

ORGANISM Drosophila melanogaster  
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

REFERENCE 1 (bases 1 to 1221)  
 AUTHORS Stapleton,M., Brokstein,P., Hong,L., Agbayani,A., Carlson,J.,  
 Champe,M., Chavez,C., Dorsett,V., Dresnek,D., Farfan,D., Frise,E.,  
 George,R., Gonzalez,M., Guarin,H., Krommiller,B., Li,P., Liao,G.,  
 Miranda,A., Mungall,C.J., Nunoo,J., Pacleib,J., Paragas,V., Park,S.,  
 Patel,S., Phouanavong,S., Wan,K., Yu,C., Lewis,S.E., Rubin,G.M.  
 and Celniker,S.

TITLE Direct Submission  
 JOURNAL Submitted (07-JUN-2002) Berkeley Drosophila Genome Project,  
 Lawrence Berkeley National Laboratory, One Cyclotron Road,  
 Berkeley, CA 94720, USA

COMMENT Sequence submitted by:  
 Berkeley Drosophila Genome Project  
 Lawrence Berkeley National Laboratory  
 Berkeley, CA 94720

This clone was sequenced as part of a high-throughput process to  
 sequence clones from Drosophila Gene Collection 1 (Rubin et al.,  
 Science 2000). The sequence has been subjected to integrity checks  
 for sequence accuracy, presence of a polyA tail and contiguity  
 within 100 kb in the genome. Thus we believe the sequence to  
 reflect accurately this particular cDNA clone. However, there are  
 artifacts associated with the generation of cDNA clones that may  
 have not been detected in our initial analyses such as internal  
 priming, priming from contaminating genomic DNA, retained introns  
 due to reverse transcription of unspliced precursor RNAs, and  
 reverse transcriptase errors that result in single base changes.  
 For further information about this sequence, including its location  
 and relationship to other sequences, please visit our Web site  
 (http://fruitfly.berkeley.edu) or send email to  
 cdna@fruitfly.berkeley.edu.

FEATURES  
 source  
 1..1221  
 /organism="Drosophila melanogaster"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:7227"

gene  
 1..1221  
 /gene="CG12919"  
 /db\_xref="FLYBASE:FBgn0033483"

CDS  
 170..955  
 /gene="CG12919"  
 /note="Longest ORF"  
 /codon\_start=1  
 /product="SD18286p"  
 /protein\_id="AAM51093.1"  
 /db\_xref="GI:21430830"  
 /db\_xref="FLYBASE:FBgn0033483"  
 /translation="MLNKNJNAHTGTTPTSETTAAGGETDSASSASNDNDVDDFTS  
 YNAHKKKQERKRSIADVRNEQNIQGNHTLEQKSSNEATSKESPAPLHRRMRHSR

## ORIGIN

Alignment Scores:  
 Pred. No.: 4,78e-105 Length: 1221  
 Score: 1626.00 Matches: 309  
 Percent Similarity: 97.78% Conservative: 0  
 Best Local Similarity: 97.78% Mismatches: 1  
 Query Match: 75.21% Indels: 7  
 DB: 3 Gaps: 1

US-09-813-329-6 (1-409) x AY119233 (1-1221)

QY 100 TTTCTGluAsnAlaLeuIleAAspTyrProLysLysValAspGlyLeuThrAspGluGluAsp 119  
 Db 6 TACGAGAATGCCCTCATCGACTATCCCAA-AAGGTGGATGGCCTCAGCGATGAGGAGGAC 64  
 QY 120 AspAspAspGlyAAspGlyLeuAspSerIleAlaAspAspGluAspAspAspValSerTyr 139  
 Db 65 GACGACCATGGCGATGGTCTGGATTCCATTGGCGACGACGAGGACGACGAGCTTAGCTAT 124  
 QY 140 SerSerValAspAspValGlyAlaAspTyrGluAspTyrThrAspMetLeuAsnLysLeu 159  
 Db 125 AGCTCTGTGGATGATGTGGCGACAGACTACGAGGACTACACCGATATGTTAAATAAACTC 184  
 QY 160 AsnAsnAlaHisThrGlyThrThrProThrSerGluThrThrAlaGluGlyGluGlyGlu 179  
 Db 185 AACAAATCCATATCCCGCACACCGCCACACTCTGAGACCACTGCTGAGGGCGAGGGCGAG 244  
 QY 180 ThrAspSerAlaSerSerAlaSerAsnAspAsnValPheAspAspPheThrSerTyr 199  
 Db 245 ACGGACAGTGCATCTCTCAGGCTCAATGATGACNATGTGTTGATGACTTTTACGAGCTAC 304  
 QY 200 AsnAlaHisLysLysLysGlnGluArgLysSerArgSerIleAlaAspValArgAsnGlu 219  
 Db 305 AATGCCCAAAAAGAGAGGAGAGAAATCTCGCTCGATTGCCGATGTACGCAATGAG 364  
 QY 220 GluGlnAsnIleGlnGlyAsnHisThrGluLeuGlnLysSerSerAsnGluAlaThr 239  
 Db 365 GAGCAGAAATATTCAGGAAATCACACAGAGCTTCAGGAAAGTCAATCCAATGAGGCAACT 424  
 QY 240 SerLysGluSerProAlaProLeuHisHisArgArgArgMetHisSerArgHisArgHis 259  
 Db 425 TCCAAAGAGAGCCCTGCACCACTTCACCACTCGCAGAGATGCATTCCTCCGCCATCGCAC 484  
 QY 260 LeuLeuValArgLys-----AlaArgSerGluAspSerArgProAla 273  
 Db 485 CTCTAGTCGCAAAAGGTGAATCTCTTTTTCAGCCAGATCCGAGGACTCGAGGCCAGCA 544  
 QY 274 AlaHisPheHisLeuSerSerArgArgArgHisGlnGlySerMetGlyTyrHisGlyAsp 293  
 Db 545 GCCCATTTCCACCTGACGACGCGCGCGTCACCAAGAAAGTATGGGCTACCATGAGAGAT 604  
 QY 294 MetTyrIleGlyAsnAspAsnGluArgAsnSerTyrGlnGlyHisPheGlnThrArgAsp 313  
 Db 605 ATGTACATAGAAATGATACGAGAGAACTCTTTATCAGGGACACTTTTCAACCGCGCAT 664  
 QY 314 GlyValLeuThrValThrAsnThrGlyLeuTyrTyrValTyrAlaGlnIleCysTyrAsn 333  
 Db 665 GGGCTCTTGACGGTGCACCAATACAGGCTATATACGTATATACGCCACAGATATGCTCAAC 724  
 QY 334 AsnSerHisAspGlnAsnGlyPheIleValPheGlnGlyAspThrProPheLeuGlnCys 353  
 Db 725 AACTCGCACACACGACGACGATTTATGCTCTTTCAAGGAGACACTCCCATTTCTCTGACATGC 784  
 QY 354 LeuAsnThrValProThrAsnMetProHisLysValHisThrCysHisThrSerGlyLeu 373  
 Db 785 TTGAACACGGTGGCCCAACATGCCCATTAAGGTGCACACCTTGCACACAGAGTGTCTG 844  
 QY 374 IleHisLeuGluArgAsnGluArgIleHisLeuLysAspIleHisAsnAspArgAsnAla 393

HRLLVRKGBESLSARSSESRPAHFHLSRRRRHQSGMGYHGDWYIIGNDNRNYSQGH  
 FQTRDGLVTVNTGLYVYVYAOICVNNSHDQNGFTVPOQDTPFLQCLNTVPTNMPHKVH  
 TCHTSGLIHLERNERIHLDKIHNDRNAVLREGNRSYFGIFKV

```

Db      845 ATCCACCTGGAACGAAACGAGGATCCATCTGAAGGACATTCACACGATCGCAATGCA 904
Qy      394 ValLeuArgGluGlyAsnAsnArgSerTyrPheGlyIlePheLysVal 409
      |||||
Db      905 GTTCTGCGGAGGAAACACCAAGGACTACTTTGGCATCTTCAAGGTG 952

RESULT 7
LOCUS   AC012854                      55359 bp      DNA      linear      HTG 03-NOV-1999
DEFINITION Drosophila melanogaster, *** SEQUENCING IN PROGRESS ***.
ACCESSION AC012854
VERSION   AC012854.1 GI:6223467
KEYWORDS HTG; HTGS, PHASE2.
SOURCE   Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
          Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
          Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
          Ephydroidea; Drosophilidae; Drosophila.
          1 (bases 1 to 55359)
          Adams,M. and Venter,J.C.
REFERENCE Direct Submission
AUTHORS   Submitted (03-NOV-1999) Celera Genomics, 45 West Gude Drive,
TITLE     Rockville, MD, USA
JOURNAL
COMMENT   This sequence was identified as CDM:10210834 by the submitter.
          For further information on this sequence you may e-mail to
          fly@celera.com.
          * NOTE: This is a 'working draft' sequence.
          * This sequence will be replaced
          * by the finished sequence as soon as it is available and
          * the accession number will be preserved.

FEATURES
source   Location/Qualifiers
          1..55359
            /organism="Drosophila melanogaster"
            /mol_type="genomic DNA"
            /db_xref="taxon:7227"

ORIGIN
Alignment Scores:
Pred. No.:      6,2e-95      Length:      55359
Score:          1508.50      Matches:    351
Percent Similarity: 43.01%      Conserva:    0
Best Local Similarity: 43.01%      Mismatches: 5
Query Match:    69.77%      Indels:    463
DB:             2           Gaps:      5

US-09-813-329-6 (1-409) x AC012854 (1-55359)

Qy      54 LeuAlaLeuThrIleTTrpGlnThrThrArgValSerHisLeuAspLysGluLeuLysSer 73
Db      31678 CTTATCTTGCAAGATCTGGCAGACAAACGCGTGTATCGCATCTGGACAAAGGAGCTGAAGAGC 31737

Qy      74 LeuLysArgValValAspAsnLeuGlnArgLeuGlyIleAsnTyrLeuAspGluPhe 93
Db      31738 CTGAAGCGAGTCGTGATATCTCCACAGCGTTGGCATAACTATCTGGACGAGTTC 31797

Qy      94 AspGluPheGlnLysGlu----- 99
Db      31798 GACGAGTTCGAAAGGAGGTGAGCAATTGGTAAATCACTTTCGAGCATCAATATTGTA 31857

Qy      99 ----- 99
Db      31858 CTCGGCGATGGGAGGGGTGTAACATTTGCCCGGGCAAAATACAGATACCTCGAATTG 31917

Qy      99 ----- 99
Db      31918 TTTACAAAGTCGGAATGATTGTTATTCCTTTGAGGTGGTACTTGTAGTTTTTAATGAGGG 31977

Qy      99 ----- 99
Db      31978 TATATCGAACTTTTGTATTGAGTATCTATTATGTTTAGTTTTCTTTTGGCTTAAGGTT 32037

Qy      99 ----- 99

```

```

Db      32038 AATTGCAATTTATGATGGCAAACTCTGGCTAGGGTATCTCACAGTCGGTAAGCCATGCG 32097
Qy      99 ----- 99
Db      32098 CTTGTTTTTGA AAAAGATGCTCTCAAAAATAATGACAGAGCTAATAATATGTTGTTTGTGT 32157
Qy      99 ----- 99
Db      32158 GGGACATACAAAATATTTATAGAAATTATGTGCTAAATAATAATTTTCGAGCGCTGTTTCAT 32217
Qy      99 ----- 99
Db      32218 CTGTTATTGTTTCAATTTTTTTTTTTGTTAGCTGCTGGCCAAAGGCACATAGAAAATTTGTG 32277
Qy      99 ----- 99
Db      32278 ATAGATTATTGAATTATAATGTCCTTTCCGCAAGGAACACTCATTTCAAATATGAGG 32337
Qy      99 ----- 99
Db      32338 CGCCAGCCAGTTACGTAATTTTATGAGATACAAATCCCAAGTGACAAAAGGTAAAAAGCC 32397
Qy      99 ----- 99
Db      32398 CCATTAATTCACAAAAATACTAGGCTCTATCAGTAGATATAGTCGTTCAAGCGCAGCGA 32457
Qy      99 ----- 99
Db      32458 TCGCTTTGGCTGTATGTGACGATTGATCGAATCTCATCAAAATGCGCTAATGCGCAGAGA 32517
Qy      99 ----- 99
Db      32518 TGGTTTTAAGATTTTAAGATTTTATTATATTTTGTATGCTGGGCACTCTGTGGTGCCAATT 32577
Qy      99 ----- 99
Db      32578 CTCTTGAGAGTTTCCTTCGCTTTTAAATATATGATGATGTACACAAGTATATGTATAGC 32637
Qy      99 ----- 99
Db      32638 AGTATCATGTTTTTTGAGCGAGCGGCATCAATTAGGATGATGACGAAACAGCCTGTGCTG 32697
Qy      99 ----- 99
Db      32698 CGGGCGTGTCTCTAAAAATGAACGTCCCAATGGATGTGAGATATGTGCCAGGACTGTGCG 32757
Qy      99 ----- 99
Db      32758 CAGCTGCAGTCACGTACTTGGCCAAAGTCTTCTGACCCAAAATGCTGTATGCCATCCCTTC 32817
Qy      100 -----TyrGluAsnAlaLeuIleAspTyrProLysLysValAspGlyLeuThrAspG 117
Db      32818 GTTGCAATACGAGAAATGCCCTCATCGACTATCCAAAAAGGTGGATGGCCTCAGCGATGA 32877
Qy      117 uGluAspAspAspGlyAspGlyLeuAspSerIleAlaAspAspGluAspAspAspVa 137
Db      32878 GGAGGACGACGACGATGCGGATGCTGGAATCCATTGCGGACGACGAGGACGACGCT 32937
Qy      137 lSerTyrSerSerValAspAspValGlyAlaAspTyrGluAspTyrThrAspMetLeuAs 157
Db      32938 TAGCTATAGCTCTGTGGATGATGTTGGCGCAGACTACGAGGACTACACCGATATGTTAAA 32997
Qy      157 nLysLeuAsnAsnAlaHisThrGlyThrThrProThrSerGluThrThrAlaGluGlyG 177
Db      32998 TAACTCAACAAATGCACATACCGCACCCACCCACACTCTGAGACCCTGCTGAGGGGGA 33057
Qy      177 uGlyGluThrAspSerAlaSerSerAlaSerAsnAspAspValPheAspAspPheTh 197
Db      33058 GGGCGAGACGACAGTGATCTCTCAGCTCAATGATGACAATGTGTTCGATGACTTTAC 33117
Qy      197 rSerTyrAsnAlaHisLysLysLysGlnGlu----- 207
Db      33118 CAGCTACAATCCCAAAAAAGAGACAGGA -GAGGTGAGCAGGATGAAATTTTTCAGGGA 33176

```



QY 74 LeuLysArgValValAspAsnLeuGlnArgLeuGlyIleAsnTyrLeuAspGluPhe 93  
Db 131484 CTGAGCGAGTCTCGAATATCTCCAGCAGCGTTTGGGCATAAACTATCTGACGAGTTC 131425  
QY 94 AspGluPheGlnLysGlu----- 99  
Db 131424 GACGAGTTCCAAAGGAGGTGAGCAATTCGGTAATCACTTCGAGCATCAATATTGTGA 131365  
QY 99 ----- 99  
Db 131364 CTCGGCGATGGGCGAGGGCTGTAACTTTGCCCGGCAAAATACAGATACCTCGAATG 131305  
QY 99 ----- 99  
Db 131304 TTTACAAAGTCGGAATGATTGTTATCTCTGAGGCTGTACTTGAGTTTAAATGAGGG 131245  
QY 99 ----- 99  
Db 131244 TATATCGAACTTTTGTATTGAGTATCTATTATGTTAGTTTCTTTTTCGCTAAAGGGT 131185  
QY 99 ----- 99  
Db 131184 AATTTGCATTTTATGATGCAAACTCTGGCTAGGGTATCTCACAGTCGGTAAGCCATGCG 131125  
QY 99 ----- 99  
Db 131124 CCTGTTTTGAAAAGAAATGCCTCAAAAATGACCAGCAGTAATATGTGTTATTGTGT 131065  
QY 99 ----- 99  
Db 131064 GGGACATACAAAATATTATAGAAATTATGTCTAAATATAAATTCGACGCTGTTTCAT 131005  
QY 99 ----- 99  
Db 131004 CTGTTATTGTTTCATTTTCTTTGTTAGCTGTTGGCCAAAGGCAATAGAAATTTGTG 130945  
QY 99 ----- 99  
Db 130944 ATAGATTATTGAATTATAATGTCTCTTCGCAAGGAACACTCATTTCAATATAGG 130885  
QY 99 ----- 99  
Db 130884 CGCAGCCAGTTCAGTAAATTATGAGATACAAATCCCAAGTGACAAAAGGTAAAGCC 130825  
QY 99 ----- 99  
Db 130824 CCAATAATTCACAAAAATACTAGGCTCTATCAGTAGATATAGTCGTTCAAGCGGAGCGA 130765  
QY 99 ----- 99  
Db 130764 TCGCTTTTGGCTGTATGTGACGATTGATCGAATCTCATCAATGGCCTAATGGCCAGAGA 130705  
QY 99 ----- 99  
Db 130704 TGGTTTTAAGATTTTAAGATTATTATATTTTGAAGCTGGGCATCTGTGTGGCCAATT 130645  
QY 99 ----- 99  
Db 130644 CTCTTGAGAGTTTCTCTTCGCTTTTAAATATATGTATGATGTACACAGTATATGTATAGC 130585  
QY 99 ----- 99  
Db 130584 AGTATCATGTTTTTGAAGCGGCGCCATCAATTAGGATGATGACGAAACAGCCTGTGTCTG 130525  
QY 99 ----- 99  
Db 130524 CGGCGGTGTCTTAAAAATGAACGTCCCAATGGATGTGAGATATGTGCCAGGACTGTGC 130465  
QY 99 ----- 99  
Db 130464 CAGTCGAGTCAGGTACTTGGCCAAAGTCTTCTGACCCAAAATGCTGATGCCCATCCCTTC 130405

QY 100 -----TyrGluAsnAlaLeuIleAspTyrProLysLysValAspGlyLeuThrAspGly 117  
Db 130404 GTTTCAGTACGAGAAATGCCCTCATCGACTATCCAAAAAGGTGGATGGCTCACGGATGA 130345  
QY 117 uGluAspAspAspGlyAspGlyLeuAspSerIleAlaAspAspGluAspAspVa 137  
Db 130344 GAGGAGCAGCAGATGGCGATGGTCTGGATTCATTTCGGGACGACGAGCAGCAGACGT 130285  
QY 137 lSerTyrSerSerValAspValGlyAlaAspTyrGluAspTyrThrAspMetLeuAs 157  
Db 130284 TAGCTATAGTCTGTGGATGATGTTGGCGCAGACTACGAGGACTACCCGATATGTTAAA 130225  
QY 157 nLysLeuAsnAsnAlaHisThrThrThrProThrSerGluThrThrAlaGluGlyGly 177  
Db 130224 TAAACTCAACAATGACATACCGGCACACCGCCACCATCTGAGACCATGTCTGAGGCGA 130165  
QY 177 uGlyGluThrAspSerAlaSerSerAlaSerAsnAspAspAsnValPheAspAspPheTh 197  
Db 130164 GGGCGAGACGGACAGTGCATCTCAGCCTCAATGATGACAAATGTGTTTCGATGACTTTAC 130105  
QY 197 rSerTyrAsnAlaHisLysLysGlnGlu----- 207  
Db 130104 CAGCTACAATGCCACAAAAGAGCAGGA-GAGGTGAGCACGGATGAAATTTTCAGGGA 130046  
QY 208 -----ArgLysSerArgSerIleAlaAspVa 216  
Db 130045 TTTATCGTATTAAAGCCTCTGCTTTTCTCTTCACAGAAAATCTCGCTCGATTGCCGATGT 129986  
QY 216 lArgAsnGluGluGlnAsnIleGlnGlyAsnHisThrGluLeuGlnGlyLysSerSerAs 236  
Db 129985 ACGCAATGAGGAGCAGATATTCAAGGAAATCACACAGAGCTTCAGGAAAAGTCTATCAA 129926  
QY 236 nGluAlaThrSerLysGlu----- 242  
Db 129925 TGAGGCAACTTCCAAAAGAGAGGTAAGTAGACGTTCCGTTGAAAAACAGCAAGAAAACC 129866  
QY 242 ----- 242  
Db 129865 ATTAATGAAGTTATATGCTATGGGTTAAAGATAGATAGTTAACTGCGTAACGTAACACA 129806  
QY 243 -----SerProAlaProLeuHisHisArgArgArgMeth 254  
Db 129805 TATGTATACATTTTCTTTTACTTTGAAGCCCTGCGACCACTTCACCACCGTCGCAATATG 129746  
QY 254 lSerArgHisArgHisLeuLeuValArgLysAla----- 265  
Db 129745 ATTCCCGCATCGCCACCTCCTAGTCGCAAAAGG-TGTGTAGTGTAGAAAAAGCGAAACAA 129687  
QY 265 ----- 265  
Db 129686 TTGGCAAAAATCGAATACCAATCTCTATTCCGACCTCTGTTAATTTGCGAGGTGAATCTC 129627  
QY 266 -----ArgSerGluAspSerArgProAlaAlaHisPheHisLeuSerSerArgA 282  
Db 129626 TTTCTTCAGCCAGATCCGAGGACTCGAGGCGCAGCAGCCCAATTTCCACTTGAGCAGCAGGC 129567  
QY 282 rGArgHisGlnGly----- 286  
Db 129566 GCGGTCCACCAAGGAAGTATGGTAAAGAGATGCTTCGCCGCTCAAGCGTGTACATACAAGT 129507  
QY 287 -----SerMetGlyTyrHisGlyAspMetT 295  
Db 129506 ACTACAGCAACCTAATGAAGTCACTTCTTGAATTCTCTCA-GGCTACCATGGAGATATGT 129448  
QY 295 rTlleGlyAsnAspAsnGluArgAsnSerTyrGlnGlyHisPheGlnThrArgAspGlyV 315  
Db 129447 ACATAGAAATGATAACGAGAGAAACTCTTATCAGGGACACTTTCAAACCGCGATGGCG 129388  
QY 315 alLeuThrValThrAsnThrGlyLeuTyrTyrValTyrAlaGlnIleCysTyrAsnAsn 335  
Db 129387 TCTTGACGGTGACCAATACAGGCTATATTACGTATACGCCCCAGATATGCTCAACAAC 129328  
QY 335 erHisAspGlnAsnGlyPheIleValPheGlnGlyAspThrProPheLeuGlnCysLeuA 355

```

Db 129327 CGCAGCAGCAGGATTTATCGCTTTTAAAGGAGACACTCCATTCCTGCGAGTGTGA 129268
Qy 355 ntThrValProThrAsnMetProHisLysValHisThrCysHisThrSerGlyLeuIleH 375
Db 129267 ACAGGTGCCACCAACATGACATAGGTGACACCTGCGACAGAGTGGTCTGATCC 129208
Qy 375 isLeuGluArgAsnGluArgIleHisLeuLysAspIleHisAsnAspArgAsnAlaValL 395
Db 129207 ACCTGGAACGAAACGAGAGGATCCATCTGAAGGACATTCACACGATCGCAATGCAGTTC 129148
Qy 395 euArgGluGlyAsnAsnArgSerTyrPheGlyIlePheLysVal 409
Db 129147 TSCGGGAGGAAACAAACCGAAGCTACTTTGGCATCTTCAAGGTG 129104

RESULT 9
AC007414
LOCUS AC007414 172904 bp DNA linear INV 06-SEP-2001
DEFINITION Drosophila melanogaster, chromosome 2R, region 46C-46D, BAC clone
          BACR06J01, complete sequence.
ACCESSION AC007414
VERSION AC007414.6 GI:15451491
KEYWORDS HTG.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
          Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
          Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
          Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 172904)
          Celnikier,S.E., Adams,M.D., Kronmiller,B., Tyler,D., Wan,K.H.,
          Holt,R.A., Evans,C.A., Gocayne,J.D., Amanatides,P.G., Brandon,R.C.,
          Rogers,Y., An,H., Baldwin,D., Banzon,J., Beeson,K.Y., Busam,D.A.,
          Carlson,J.W., Center,A., Champe,M., Davenport,L.B., Dietz,S.M.,
          Dodson,K., Dorsett,V., Doup,L.E., Doyle,C., Dresnek,D., Farfan,D.,
          Ferrera,S., Frise,E., Galle,R.F., Garg,N.S., George,R.A.,
          Gonzalez,M., Houck,J., Hoskins,R.A., Hostin,D., Howland,T.J.,
          Ibegwan,C., Jalali,M., Kruse,D., Li,P., Mattei,B., Moshrefi,A.,
          McIntosh,T.C., Moy,M., Murphy,B., Nelson,C., Nelson,K.A., Nunoo,J.,
          Pacleb,J., Paragas,V., Park,S., Patel,S., Pfeiffer,B.,
          Phouanavong,S., Pittman,G.S., Puri,V., Richards,S., Scheeler,F.,
          Stapleton,M., Strong,R., Svirskaas,R., Tector,C., Williams,S.M.,
          Zaveri,J.S., Smith,H.O., Rubin,G.M. and Venter,J.C.
          Sequencing of Drosophila chromosome 2R, region 46C-46D
          Unpublished
2 (bases 1 to 172904)
          Celnikier,S.E., Agbavani,A., Arcaina,T.T., Baxter,E., Blazej,R.G.,
          Butenhoff,C., Champe,M., Chavez,C., Chew,M., Ciesiolka,L.,
          Doyle,C.M., Farfan,D.E., Galle,R., George,R.A., Harris,N.L.,
          Hoskins,R.A., Houston,K.A., Hummasti,S.R., Karra,K., Kearney,L.,
          Kim,E., Lee,B., Lewis,S., Li,P., Lomotan,M.A., Mazda,P.,
          Moshrefi,A.R., Moshrefi,M., Nixon,K., Pacleb,J.M., Park,S.,
          Pfeiffer,B., Poon,L., Sequeira,A., Sethi,H., Snir,E.,
          Svirskaas,R.R., Wan,K.H., Weinburg,T., Zhang,R., Zieran,L.L. and
          Rubin,G.M.
          Direct Submission
          Submitted (28-APR-1999) Drosophila Genome Center, Lawrence Berkeley
          Laboratory, MS 64-121, Berkeley, CA 94720, USA
          On Sep 6, 2001 this sequence version replaced gi:13324748.
          Sequence submitted by:
          Lawrence Berkeley National Laboratory, MS 64-121
          Berkeley, CA 94720
          This sequence was assembled using end sequences from a whole genome
          shotgun and from subclones of this BAC and its neighboring clones.
          For further information about this sequence, including its location
          and relationship to other sequences, please visit our sequence
          archive Web site (http://www.fruitfly.org/sequence/) or send email
          to bdgp@fruitfly.berkeley.edu.
          Location/Qualifiers
FEATURES             source
1..172904
   /organism="Drosophila melanogaster"
   /mol_type="genomic DNA"
   /strain="y; cn bw sp"

```

```

/db_xref="taxon:7227"
/chromosome="2R"
/map="46C-46D"
/clone="BACR06J01 (D601)"
/clone_lib="RPCI-98 (Roswell Park Cancer Institute
Drosophila melanogaster BAC library, partial EcoRI in
pBACe3.6)"

ORIGIN
Alignment Scores:
Pred. No.:      2.27e-94      Length:      172904
Score:          1508.50      Matches:      351
Percent Similarity: 43.01%      Conservative: 0
Best Local Similarity: 43.01%      Mismatches: 5
Query Match:     69.77%      Indels:      463
DB:              3          Gaps:         5

US-09-813-329-6 (1-409) x AC007414 (1-172904)
Qy 54 LeuAlaLeuThrIleTrpGlnThrThrArgValSerHisLeuApsLysGluLeuLysSer 73
Db 136338 CTTATCTTGCAGATCTGGCAGACACACGCGTGTATCGCATCTGGACAGGAGCTCAAGAGC 136397
Qy 74 LeuLysArgValValAspAsnLeuGlnArgLeuGlyIleAsnTyrLeuApsGluPhe 93
Db 136398 CTGAAGCGAGTCGTCGATAATCTCCAGCAGCGTTTGGGCATAAACTATCTGGACGAGTTC 136457
Qy 94 AspGluPheGlnLysGlu----- 99
Db 136458 GACGAGTTCCTCAAAAGGAGGTGAGCAATTCGCTAAATCACTTTTCGAGCATCAATATTGTA 136517
Qy 99 ----- 99
Db 136518 CTGGCGATGGGCGAGGGCTGTAAACATTTGCCCGGGCAATACAGATACCTCGAATTG 136577
Qy 99 ----- 99
Db 136578 TTTACAAAGTCGGAATGATTGTTATTCCTTGAGGCTGTACTTGAGTTTTTAATGAGGG 136637
Qy 99 ----- 99
Db 136638 TATATCGAACTTTTGTATTATGAGTATCTATTATGTTAGTTTCTTTTTCCTAAAGGGT 136697
Qy 99 ----- 99
Db 136698 AATTGCAATTTATGATGGCAAACTCTGGCTAGGGTATCTCACAGTCGGTAAGCCATGCG 136757
Qy 99 ----- 99
Db 136758 CCTGTTTTTGAAAAAGAAATGCCTCAAAAAAATGACCAGACGTAATATGTGTTATTGTGT 136817
Qy 99 ----- 99
Db 136818 GGGACATACAAAAATATTTATAGAAATTATGTGCTAAATAATAATTTTCGACGCTGTTTCAT 136877
Qy 99 ----- 99
Db 136878 CTGTTATTGTTTCATTTTTTTTTTTTGTAGCTGCTTGCCAAAGGCACATAGAAAAATTGTG 136937
Qy 99 ----- 99
Db 136938 ATAGATTATTGTAATTATAATGTCCTTTCCGCAAGAAACACTCATTTCAAATATGAGG 136997
Qy 99 ----- 99
Db 136998 CGCCAGCAGTTCACGCTAAATTTATGAGATACAAATCCCCAAGTGACAAAAAGGTAAGGCC 137057
Qy 99 ----- 99
Db 137058 CCAATAATTCAAAAAAATACTAGGCTCTATCAGTAGATATAGTCGTTCAAGGCGACGA 137117
Qy 99 ----- 99

```



Gonzalez,M., Houck,J., Hoskins,R.A., Hostin,D., Hostin,D., Howland,T.J., Ibegwan,C., Jalali,M., Kruse,D., Li,P., Mattei,B., Moshrefi,A., McIntosh,T.C., Moy,M., Murphy,B., Nelson,C., Nelson,K.A., Nunoo,J., Pacleb,J., Paragas,V., Park,S., Patel,S., Pfeiffer,B., Phouanavong,S., Pittman,G.S., Puri,V., Richards,S., Scheeler,F., Stapleton,M., Strong,R., Svirskaas,R., Tector,C., Williams,S.M., Zaveri,J.S., Smith,H.O., Rubin,G.M. and Venter,J.C.

## Direct Submission

Submitted (08-NOV-2001) Berkeley Drosophila Genome Project, MS 64-121, Lawrence Berkeley National Laboratory, One Cyclotron Road, Berkeley, CA 94720, US

## Sequence submitted by:

Berkeley Drosophila Genome Project  
Lawrence Berkeley National Laboratory, MS 64-121  
Berkeley, CA 94720

This sequence was assembled using end sequences from a whole genome shotgun and from subclones of this BAC and its neighboring clones. For further information about this sequence, including its location and relationship to other sequences, please visit our sequence archive web site (<http://www.fruitfly.org/sequence/>) or send email to [bdgp@fruitfly.berkeley.edu](mailto:bdgp@fruitfly.berkeley.edu).

## FEATURES

Location/Qualifiers

source

1..189620  
/organism="Drosophila melanogaster"  
/mol\_type="genomic DNA"  
/strains="y; cn bw sp"  
/db\_xref="taxon:7227"  
/chromosome="2R"  
/map="46D-47A"  
/clone="BACR09J12 (D1345)"  
/clone\_lib="RPCI-98 (Roswell Park Cancer Institute  
Drosophila melanogaster BAC library, partial EcoRI in  
pBACe3.6)"

## ORIGIN

## Alignment Scores:

Pred. No.: 2.52e-94 Length: 189620  
Score: 1508.50 Matches: 351  
Percent Similarity: 43.01% Conservative: 0  
Best Local Similarity: 43.01% Mismatches: 5  
Query Match: 69.77% Indels: 463  
DB: 3 Gaps: 5

US-09-813-329-6 (1-409) x AC099029 (1-189620)

Qy 54 LeuAlaLeuThrIleTrrGlnThrThrArgValSerHisLeuAspLysGluLeuLysSer 73  
Db 166400 CTTATCTTGCAGATCTGCAGACAAACGCGTGTATCGCATCTGGACAGGAGCTGAAGAGC 166341  
Qy 74 LeuLysArgValValAspAsnLeuGlnGlnArgLeuGlyIleAsnTyrLeuAspGluPhe 93  
Db 166340 CTGAAGCGAGTCGTGATAAATCTCAGCAGCGTTTGGGCATAAACTATCTGGACGAGTTC 166281  
Qy 94 AspGluPheGlnLysGlu----- 99  
Db 166280 GACGAGTTCCAAAGAGGCTGAGCAATTGCGTAATATCACTTTCGAGCATCAATATTTGTA 166221  
Qy 99 ----- 99  
Db 166220 CTCGGCGATGGGCGAGGGCTGTAAATTTGCCCCGGGCAAAATACAGATACCTCGAATG 166161  
Qy 99 ----- 99  
Db 166160 TTTACAAAGTCGGAATGATTGTTATTCCTTTGAGGCTGGTACTTGAGTTTTTAATGAGGG 166101  
Qy 99 ----- 99  
Db 166100 TATATCGAACTTTTGTATTGAGTATCTATTATGTTAGTTTCTTTTGGCTAAAGGGT 166041  
Qy 99 ----- 99  
Db 166040 AATTTGCAATTTTATGATGCGAAACTCTGGCTAGGGTATCTCACAGTCGGTAAGCCATCGC 165981

Qy 99 ----- 99  
Db 165980 CCTGTTTTGAAAAAGAAATGCCTCAAAAAAATGACCAGACGTAATATGTGTTATTGTGT 165921  
Qy 99 ----- 99  
Db 165920 GGGACATACAAAAATATTTATAGAATTTATGTCTAAAAATATAAATTTGACGCGTGTTCAT 165861  
Qy 99 ----- 99  
Db 165860 CTGTTATTGTTTCATTTTTTTTTTTGTTAGCTGCTTGCCCAAGGACACATAGAAAAATTGTG 165801  
Qy 99 ----- 99  
Db 165800 ATAGATTATTGAATTATAATGTCTCTTCGCAAGGAACACACTCAATTTCAAATATGAGG 165741  
Qy 99 ----- 99  
Db 165740 CGCCAGCGAGTTCACGTAATTTATGAGATACAAATCCCAAGTGACAAAAGGTAAAGCC 165681  
Qy 99 ----- 99  
Db 165680 CCAATAATTCAAAAAAATACTAGGCTCTATCAGTAGATATAGTCGTTCAAGGCGGAGCGA 165621  
Qy 99 ----- 99  
Db 165620 TCGCTTTTGGCTGTATGTGACGATTGATCGAATCTCATCAATGGCCTTAATGGCCAGAGA 165561  
Qy 99 ----- 99  
Db 165560 TGGTTTTAAGATTTTAAGATTTATTTATATTTTATGATGGCTGGGCATCTGTGGTGCCTCAAT 165501  
Qy 99 ----- 99  
Db 165500 CTCTTGAGAGTTTCCTTCGCTTTTAAATATATGTATGATGTACACAAGTATATGTATACG 165441  
Qy 99 ----- 99  
Db 165440 AGTATCATGTTTTTTGAGCGAGGCGCCATCAATTAGGATGATGACGAAACAGCCTGTGCTG 165381  
Qy 99 ----- 99  
Db 165380 CGGCGGTGTGTCTAAAAAATGAACGTCGCCAAATGGATGTGAGATATGTGCCAGGACTGTGC 165321  
Qy 99 ----- 99  
Db 165320 CAGCTGCAGTCACGTACTTGGCCAAAGTCTTCTGACCCAAATGCTGATGCCCATCCCTTC 165261  
Qy 100 -----TyrGluAsnAlaLeuIleAspTyrProLysLysValAspGlyLeuThrAspGly 117  
Db 165260 GTTCAGTACGAGAAATGCCCTCATCGACTATCCAAAAAAGGTGGATGGCCTCAGCGATGA 165201  
Qy 117 uGluAspAspAspGlyAspGlyLeuAspSerIleAlaAspAspGluAspAspAspVa 137  
Db 165200 GGAGGACGACGACGATCGCATGCTGTGATTCATTTGCCGACGACGAGGACGACGAGT 165141  
Qy 137 lSerTyrSerSerValAspValGlyAlaAspTyrGluAspTyrThrAspMetLeuAs 157  
Db 165140 TAGCTATAGCTCTGTGGATGATGTTGGCGCAGACTACGAGGACTACACCGATATGTTAAA 165081  
Qy 157 nLysLeuAsnAsnAlaHisThrGlyThrThrProThrSerGluThrThrAlaGluGlyGly 177  
Db 165080 TAAACTCAAAATGCACATACCGGCACACGCCCCACATCTGAGACCACCTGCTGAGGCGCA 165021  
Qy 177 uGlyGluThrAspSerAlaSerSerAlaSerAsnAspAspAsnValPheAspAspPheTh 197  
Db 165020 GGGCGAGACGACAGTGCATCTCAGCCTCAAAATGATGACAAATGTGTTCGATGACTTTAC 164961  
Qy 197 rSerTyrAsnAlaHisLysLysLysGlnGlu----- 207  
Db 164960 CAGCTACATGCCCAAAAGAAAGACAGGA- GAGGTGAGCAGCGATGAAATTTTCAGGGA 164902  
Qy 208 -----ArgLysSerArgSerIleAlaAspVa 216



Db	164901	TTTATCGTATTAAAGCCTCTGCTTTTCTCTTACAGAAATCTCGCTCCGATGCGCATGT	164842
QY	216	lArGAsnGluGluGlnAsnIleGlnGlyAsnHisThrGluLeuGlnGluYSerSxAs	236
Db	164841	ACGCAATGAGGACAGCAATATTCAAGGAAATCACAGAGCTTCAGGAAAGTCATCCNA	164782
QY	236	nGluAlaThrSerLysGlu-----	242
Db	164781	TGAGGCACTTCCAAAGAGAGGTAGTAACGTTCCGTTGAAAAACAGCAAGAAAAACC	164722
QY	242	-----	242
Db	164721	ATTAAATGAAGTTATATGCTATGCTGTGGTTAAAGATAGTAGTAACTCGGTAACTGAAACACA	164662
QY	243	-----SerProAlaProLeuHisHisArgArgMeth	254
Db	164661	TATGTATACATTTCTTTTACTTGAAGCCCTGCACCACTTCACCACTGCAGATGC	164602
QY	254	isSerArgHisArgHisLeuLeuValAtqLysAla-----	265
Db	164601	ATTCCGCCCATGCCACCTCTTAGTCCGCAAGG-TGTGTAGTGTAGAAAAAGCGAAACAA	164543
QY	265	-----	265
Db	164542	TTGGCAAAATCGAATACCAATCTCTATTGCACTCTGTTAATTTGCAGGTGAATCTC	164483
QY	266	-----ArgSerGluAspSerArgProAlaAlaHisPheHisLeuSerSerArgA	282
Db	164482	TTCTTTGACCGAGATCCGAGGACTCGAGCCAGCAGCCCATTTCCACTGTGACGACGCG	164423
QY	282	tGArgHisGlnGly-----	286
Db	164422	GCGTGCACCAAGGAAGTATGGTAGGAGATGCTTCGCCGTCAACGCTGTACATACAAGT	164363
QY	287	-----SerMetGlyTyrHisGlyAspMet	295
Db	164362	ACTACAGCAACCTAATGAAGTCACTTCTGAAATTTCTCA-GGCTACCATGGAGATATGT	164304
QY	295	yrIleGlyAsnAspAsnGluArgAsnSerTyrGlnGlyHisPheGlnThrArgAspGlyV	315
Db	164303	ACATAGGAAATGATAACGAGAAACTCTTATCAGGACACTTTCAACGCGCGATGGCG	164244
QY	315	alLeuThrValThrAsnThrGlyLeuTyrValTyrAlaGlnIleCysTyrAsnAsnS	335
Db	164243	TCITGACGCTGACCAATACAGGCCTATATTACGTATACGCCAGATATGCTACACAAC	164184
QY	335	erHisAspGlnAsnGlyPheIleValPheGlnGlyAspThrProPheLeuGlnCysLeuA	355
Db	164183	CGCAGCACCAACGGATTTATCTCTTTCAAGGAGACATCCATTCCTCGAGTGTCTGA	164124
QY	355	snThrValProThrAsnMetProHisLysValHisThrCysHisThrSerGlyLeuIleH	375
Db	164123	ACACGCTGCCCAACATGCCACATAGGTGCACACTCCACAGAGTGGTCTGATCC	164064
QY	375	isLeuGluArgAsnGluArgIleHisLeuLysAspIleHisAsnAspArgAsnAlaVal	395
Db	164063	ACCTGGACCAACGAGAGATCCATCTGAAGGACATTCACAACGATCGCAATGCGATTC	164004
QY	395	euArgGluGlyAsnAsnArgSerTyrPheGlyIlePheLysVal	409
Db	164003	TGCGGGAGGGAACACCGAGCTACTTTGGCATCTTCAAGGTG	163960
RESULT 11	AE003831		
LOCUS	AB003831	275390 bp	DNA linear INV 14-FEB-2003
DEFINITION	Drosophila melanogaster chromosome 2R, section 19 of 74 of the complete sequence		
ACCESSION	AE003831	AE002787	AE013599
VERSION	AE003831.3	GI:21645483	
KEYWORDS	Drosophila melanogaster (fruit fly)		
SOURCE	Drosophila melanogaster		
ORGANISM	Drosophila melanogaster		

REFERENCE  
AUTHORS

Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

REFERENCE  
AUTHORS

1 (bases 1 to 275390)

Adams,M.D., Celniker,S.E., Holt,R.A., Evans,C.A., Gocayne,J.D., Ananidis,P.G., Scherer,S.E., Li,P.W., Hoskins,R.A., Galle,R.F., George,R.A., Lewis,S.E., Richards,S., Ashburner,M., Henderson,S.N., Sutton,G.G., Wortman,J.R., Yandell,M.D., Zhang,Q., Chen,L.X., Brandon,R.C., Rogers,J.H., Blazej,R.G., Champe,M., Pfeiffer,B.D., Wan,K.H., Doyle,C., Baxter,E.G., Helt,G., Nelson,C.R., Gabor,G.L., Abail,J.F., Agbayani,A., An,H.J., Andrews-Pfannkoch,C., Baldwin,D., Ballew,R.M., Basu,A., Baxendale,J., Bayraktaroglu,L., Beasley,E.M., Beeson,K.Y., Benos,P.V., Bertram,B.P., Bhandari,D., Bhatnagar,S., Borkov,D., Botchan,M.R., Bouck,J., Brokstein,P., Brotter,P., Burtis,K.C., Busam,D.A., Butler,H., Cadieu,E., Center,A., Chandra,I., Cherry,J.M., Cawley,S., Dahlke,C., Davenport,L.B., Davies,P., de Pablos,B., Delcher,A., Deng,Z., Mays,A.D., Dew,I., Dietz,S.M., Dodson,K., Doup,L.E., Downes,M., Dugan-Rocha,S., Dunkov,B.C., Dunn,P., Durbin,K.J., Evangelista,C.C., Ferraz,C., Ferrieres,S., Fleischmann,W., Fosler,C., Gabrielian,A.E., Garg,N.S., Gelbart,W.M., Glasser,K., Glodek,A., Gong,F., Gorrell,J.H., Gu,Z., Guan,P., Harris,M., Harris,N.L., Harvey,D., Heiman,T.J., Hernandez,J.R., Houck,J., Hostin,D., Houston,K.A., Howland,T.J., Wei,M.H., Ibegwam,C., Jalali,M., Kalush,F., Karpen,G.H., Ke,Z., Kennits,J.A., Ketchum,K.A., Kimmel,B.E., Kodira,C.D., Kraft,C., Kravitz,S., Kulp,D., Lai,Z., Lasko,P., Lei,Y., Levitsky,A., Li,J., Li,Z., Liang,X., Lin,X., Liu,X., Mattei,B., McIntosh,T.C., McLeod,M.P., McPherson,D., Merkulov,G., Milshina,N.V., Mobarry,C., Morris,J., Moshrefi,A., Mount,S.M., Moy,M., Murphy,B., Murphy,L., Muzny,D.M., Nelson,D.L., Nelson,D.R., Nelson,K.A., Nixon,K., Nuskern,D.R., Pacle,J.M., Palazzolo,M., Pittman,G.S., Pan,S., Pollard,J., Puri,V., Reese,M.G., Reinert,K., Remington,K., Sanders,R.D., Scheeler,F., Shen,H., Shue,B.C., Siden-Kiamos,I., Simpson,M., Skupski,M.P., Smith,T., Spier,E., Spradling,A.C., Stapleton,M., Strong,R., Sun,E., Svirskaas,R., Tector,C., Turner,R., Venter,E., Wang,A.H., Wang,X., Wang,Z.Y., Wasserman,D.A., Weinstock,G.M., Weissbach,J., Williams,S.M., Woodagef, Worley,K.C., Wu,D., Yang,S., Yao,Q.A., Ye,J., Yeh,F.F., Zaveri,J.S., Zhan,M., Zhang,G., Zhao,Q., Zheng,L., Zheng,X.H., Zhong,F.N., Zhong,W., Zhou,X., Zhu,S., Zhu,X., Smith,H.O., Gibbs,R.A., Myers,E.W., Rubin,G.M. and Venter,J.C.

TITLE  
MEDLINE  
PUBMED

The genome sequence of Drosophila melanogaster  
Science 287 (5461), 2185-2195 (2000)  
20156006  
10731132

REFERENCE  
AUTHORS

2 (bases 1 to 275390)  
Celniker,S.E., Adams,M.D., Kronmiller,B., Wan,K.H., Holt,R.A., Evans,C.A., Gocayne,J.D., Ananidis,P.G., Brandon,R.C., Rogers,Y., Banzon,J., An,H., Baldwin,D., Banzon,J., Beeson,K.Y., Busam,D.A., Carlson,J.W., Center,A., Champe,M., Davenport,L.B., Dietz,S.M., Dodson,K., Dorsett,V., Doup,L.E., Doyle,C., Dresnek,D., Farfan,D., Ferrieres,S., Frise,B., Galle,R.F., Garg,N.S., George,R.A., Gonzalez,M., Houck,J., Hoskins,R.A., Hostin,D., Howland,T.J., Ibegwam,C., Jalali,M., Kruse,D., Li,P., Mattei,B., Moshrefi,A., McIntosh,T.C., Moy,M., Murphy,B., Nelson,C., Nelson,K.A., Nunoo,J., Pacle,J., Paragas,V., Park,S., Patel,S., Pfeiffer,B., Scheeler,F., Shouanavong,S., Sittman,G.S., Puri,V., Richards,S., Scheeler,F., Stapleton,M., Strong,R., Svirskaas,R., Tector,C., Tyler,D., Williams,S.M., Zaveri,J.S., Smith,H.O., Venter,J.C. and Rubin,G.M.

TITLE  
JOURNAL  
REFERENCE  
AUTHORS

Sequencing of Drosophila melanogaster genome  
Unpublished  
3 (bases 1 to 275390)

Misra,S., Crosby,M.A., Matthews,B.B., Bayraktaroglu,L., Campbell,K., Hradscky,P., Huang,Y., Kaminker,J.S., Prochnik,S.E., Smith,C.D., Tupy,J.L., Bergman,C.M., Berman,B.P., Carlson,J.W., Celniker,S.E., Clamp,M.E., Drysdale,R.A., Emmert,D., Frise,E., de Grey,A.D.N.J., Harris,N.L., Kronmiller,B., Marshall,B., Millburn,G.H., Richter,J., Russo,S., Searle,S.M.J., Smith,E., Shu,S., Smutniak,F., Whitfield,E.J., Ashburner,M., Gelbart,W.M., Rubin,G.M., Mungall,C.J. and Lewis,S.E.

TITLE  
JOURNAL  
REFERENCE

Annotation of Drosophila melanogaster genome  
Unpublished  
4 (bases 1 to 275390)



AUTHORS	Adams,M.D., Celniker,S.E., Gibbs,R.A., Rubin,G.M. and Venter,C.J.		
TITLE	Direct Submission		
JOURNAL	Submitted (21-MAR-2000) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA		
REFERENCE	5 (bases 1 to 275390)		
AUTHORS	FlyBase		
CONSTRM	Direct Submission		
TITLE	Submitted (06-SEP-2002) University of California Berkeley, 539 Life Sciences Addition, Berkeley, CA 94720, USA		
JOURNAL	6 (bases 1 to 275390)		
REFERENCE	FlyBase		
CONSTRM	Direct Submission		
TITLE	Submitted (23-JAN-2003) University of California Berkeley, 539 Life Sciences Addition, Berkeley, CA 94720, USA		
JOURNAL	On Jul 1, 2002 this sequence version replaced gi:10727672.		
COMMENT	Location/Qualifiers		
FEATURES	source		
	1..275390		
	/organism="Drosophila melanogaster"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:7227"		
	/chromosome="2R"		
	/note="Genotype: y[1]; cn[1] bw[1] sp[1]; Rh6[1]"		
gene	197..2954		
	/locus_tag="CG1671"		
	/note="Last curated on Mon Feb 11 17:53:43 PST 2002"		
	/map="46B7-46B9"		
	/db_xref="FLYBASE:FBgn0033454"		
mRNA	join(197..286,346..2954)		
	/locus_tag="CG1671"		
	/product="CG1671-RA"		
	/db_xref="FLYBASE:FBgn0033454"		
CDS	join(261..286,346..2683)		
	/locus_tag="CG1671"		
	/note="CG1671 gene product"		
	/codon_start=1		
	/product="CG1671-PA"		
	/protein_id="AAF58888.1"		
	/db_xref="GI:7303841"		
	/db_xref="FLYBASE:FBgn0033454"		
	/translation="MLAGISEKSYAAEARVYFNVAGGDIAWSPNGHLYCLINGSANVQ VQTSGLNSYSSPADNKRAELDNIDVEDTITCGLSQEHVLTVHRSGLLR L WDLTGKLVKQAQKGPVIRVEFPCGRLICITSGGADTLRWDSNNCLGALKD PGPAWLLVHPNPVLRKEIYAGGSDNTYAWNYETKLVHMKRGHLQVTLGLSPKSA SDCEIIVTSVDKVMYVQLQEQLESKQLKVLPLYDLEGVVYADEQLIIVASGSK LQQVDPKSKIRDLISGSDFQISRLHCSELKQLAVTTEQNIIVYDVEITLPIKHLV GYNDIILDMCPMGSHDRYLAVATNSKHFLYDTEHDMNCKLIVGHSDTVMSLAASQNL LISVKDCSVRLWLRLQDKCSLEALTQQANCHTSTIGCVAMTNGATGFASVSQDGS MKVQLVRSKEDRNSYFNLRYYAALSHDKVCNVAYAPNNKLIATAGQDKTAKVWLAE SNTLQGVLRGTRGVSWRPSVPDQIVLTSSSDCTLRISWISNFCIKRPFDOECTILR AEFIDHGKFIISASDGILLKWNITKNTCLQSLDEHNDRVWALVARSNRFFYTGGA DSKLIRGQDVTQVTRNEALDRQAALQEQTLISLHAQQLKAFVIALNLDKPKAS FDIINHFVRKDEPLQRLVDQNLVDORVALLOHVKAWTTNSRHSQAGNMILKYLIGD ALLDPKARFYNGNVEVLTPTVQVRFKRVTELNKDLAFLEFIVKCM"		
	complement(2844..13679)		
	/gene="BcDNA:GH06348"		
gene	/locus_tag="CG1516"		
	/note="Last curated on Tue Mar 12 18:58:27 PST 2002"		
	/map="46B9-46B12"		
	/db_xref="FLYBASE:FBgn0027580"		
mRNA	complement(join(2844..3521,3590..4619,4684..5341,6647..8156,8220..8346,13587..13679))		
	/gene="BcDNA:GH06348"		
	/locus_tag="CG1516"		
	/product="CG1516-RE"		
	/db_xref="FLYBASE:FBgn0027580"		
mRNA	complement(join(2844..3521,3590..4619,4684..5341,5819..5866,6647..8156,8220..8346,12506..12893))		
	/gene="BcDNA:GH06348"		
	/locus_tag="CG1516"		
	/product="CG1516-R"		
	/db_xref="FLYBASE:FBgn0027580"		
mRNA	complement(join(2844..3521,3590..4619,4684..5341,5819..5866,6647..8156,8220..8346,12506..12893))		
	/gene="BcDNA:GH06348"		
	/locus_tag="CG1516"		
	/product="CG1516-R"		
	/db_xref="FLYBASE:FBgn0027580"		
mRNA	complement(join(2844..3521,3590..4619,4684..5341,5819..5866,6647..8156,8220..8346,12506..12893))		
	/gene="BcDNA:GH06348"		
	/locus_tag="CG1516"		
	/product="CG1516-R"		
	/db_xref="FLYBASE:FBgn0027580"		
mRNA	complement(join(2844..3521,3590..4619,4684..5341,5819..5866,6647..8156,8220..8346,12506..12893))		
	/gene="BcDNA:GH06348"		
	/locus_tag="CG1516"		
	/product="CG1516-R"		
	/db_xref="FLYBASE:FBgn0027580"		
mRNA	complement(join(2844..3521,3590..4619,4684..5341,5819..5866,6647..8156,8220..8346,12506..12893))		
	/gene="BcDNA:GH06348"		
	/locus_tag="CG1516"		
	/product="CG1516-R"		
	/db_xref="FLYBASE:FBgn0027580"		
mRNA	complement(join(2844..3521,3590..4619,4684..5341,5819..5866,6647..8156,8220..8346,12506..12893))		
	/gene="BcDNA:GH06348"		
	/locus_tag="CG1516"		
	/product="CG1516-R"		
	/db_xref="FLYBASE:FBgn0027580"		
mRNA	complement(join(2844..3521,3590..4619,4684..5341,5819..5866,6647..8156,8220..8346,12506..12893))		
	/gene="BcDNA:GH06348"		
	/locus_tag="CG1516"		
	/product="CG1516-R"		
	/db_xref="FLYBASE:FBgn0027580"		
mRNA	complement(join(2844..3521,3590..4619,4684..5341,5819..5866,6647..8156,8220..8346,12506..12893))		
	/gene="BcDNA:GH06348"		
	/locus_tag="CG1516"		
	/product="CG1516-R"		
	/db_xref="FLYBASE:FBgn0027580"		
mRNA	complement(join(2844..3521,3590..4619,4684..5341,5819..5866,6647..8156,8220..8346,12506..12893))		
	/gene="BcDNA:GH06348"		
	/locus_tag="CG1516"		
	/product="CG1516-R"		
	/db_xref="FLYBASE:FBgn0027580"		
mRNA	complement(join(2844..3521,3590..4619,4684..5341,5819..5866,6647..8156,8220..8346,12506..12893))		
	/gene="BcDNA:GH06348"		
	/locus_tag="CG1516"		
	/product="CG1516-R"		
	/db_xref="FLYBASE:FBgn0027580"		
mRNA	complement(join(2844..3521,3590..4619,4684..5341,5819..5866,6647..8156,8220..8346,12506..12893))		
	/gene="BcDNA:GH06348"		
	/locus_tag="CG1516"		
	/product="CG1516-R"		
	/db_xref="FLYBASE:FBgn0027580"		
mRNA	complement(join(2844..3521,3590..4619,4684..5341,5819..5866,6647..8156,8220..8346,12506..12893))		
	/gene="BcDNA:GH06348"		
	/locus_tag="CG1516"		
	/product="CG1516-R"		
	/db_xref="FLYBASE:FBgn0027580"		
mRNA	complement(join(2844..3521,3590..4619,4684..5341,5819..5866,6647..8156,8220..8346,12506..12893))		
	/gene="BcDNA:GH06348"		
	/locus_tag="CG1516"		
	/product="CG1516-R"		
	/db_xref="FLYBASE:FBgn0027580"		
mRNA	complement(join(2844..3521,3590..4619,4684..5341,5819..5866,6647..8156,8220..8346,12506..12893))		
	/gene="BcDNA:GH06348"		
	/locus_tag="CG1516"		
	/product="CG1516-R"		
	/db_xref="FLYBASE:FBgn0027580"		
mRNA	complement(join(2844..3521,3590..4619,4684..5341,5819..5866,6647..8156,8220..8346,12506..12893))		
	/gene="BcDNA:GH06348"		
	/locus_tag="CG1516"		
	/product="CG1516-R"		
	/db_xref="FLYBASE:FBgn0027580"		
mRNA	complement(join(2844..3521,3590..4619,4684..5341,5819..5866,6647..8156,8220..8346,12506..12893))		
	/gene="BcDNA:GH06348"		
	/locus_tag="CG1516"		
	/product="CG1516-R"		
	/db_xref="FLYBASE:FBgn0027580"		
mRNA	complement(join(2844..3521,3590..4619,4684..5341,5819..5866,6647..8156,8220..8346,12506..12893))		
	/gene="BcDNA:GH06348"		
	/locus_tag="CG1516"		
	/product="CG1516-R"		
	/db_xref="FLYBASE:FBgn0027580"		
mRNA	complement(join(2844..3521,3590..4619,4684..5341,5819..5866,6647..8156,8220..8346,12506..12893))		
	/gene="BcDNA:GH06348"		
	/locus_tag="CG1516"		
	/product="CG1516-R"		
	/db_xref="FLYBASE:FBgn0027580"		
mRNA	complement(join(2844..3521,3590..4619,4684..5341,5819..5866,6647..8156,8220..8346,12506..12893))		
	/gene="BcDNA:GH06348"		
	/locus_tag="CG1516"		
	/product="CG1516-R"		
	/db_xref="FLYBASE:FBgn0027580"		
mRNA	complement(join(2844..3521,3590..4619,4684..5341,5819..5866,6647..8156,8220..8346,12506..12893))		
	/gene="BcDNA:GH06348"		
	/locus_tag="CG1516"		
	/product="CG1516-R"		
	/db_xref="FLYBASE:FBgn0027580"		
mRNA	complement(join(2844..3521,3590..4619,4684..5341,5819..5866,6647..8156,8220..8346,12506..12893))		
	/gene="BcDNA:GH06348"		
	/locus_tag="CG1516"		
	/product="CG1516-R"		
	/db_xref="FLYBASE:FBgn0027580"		
mRNA	complement(join(2844..3521,3590..4619,4684..5341,5819..5866,6647..8156,8220..8346,12506..12893))		
	/gene="BcDNA:GH06348"		
	/locus_tag="CG1516"		
	/product="CG1516-R"		
	/db_xref="FLYBASE:FBgn0027580"		
mRNA	complement(join(2844..3521,3590..4619,4684..5341,5819..5866,6647..8156,8220..8346,12506..12893))		
	/gene="BcDNA:GH06348"		
	/locus_tag="CG1516"		
	/product="CG1516-R"		
	/db_xref="FLYBASE:FBgn0027580"		
mRNA	complement(join(2844..3521,3590..4619,4684..5341,5819..5866,6647..8156,8220..8346,12506..12893))		
	/gene="BcDNA:GH06348"		
	/locus_tag="CG1516"		
	/product="CG1516-R"		
	/db_xref="FLYBASE:FBgn0027580"		
mRNA	complement(join(2844..3521,3590..4619,4684..5341,5819..5866,6647..8156,8220..8346,12506..12893))		
	/gene="BcDNA:GH06348"		
	/locus_tag="CG1516"		
	/product="CG1516-R"		
	/db_xref="FLYBASE:FBgn0027580"		
mRNA	complement(join(2844..3521,3590..4619,4684..5341,5819..5866,6647..8156,8220..8346,12506..12893))		
	/gene="BcDNA:GH06348"		
	/locus_tag="CG1516"		
	/product="CG1516-R"		
	/db_xref="FLYBASE:FBgn0027580"		
mRNA	complement(join(2844..3521,3590..4619,4684..5341,5819..5866,6647..8156,8220..8346,12506..12893))		
	/gene="BcDNA:GH06348"		
	/locus_tag="CG1516"		
	/product="CG1516-R"		
	/db_xref="FLYBASE:FBgn0027580"		
mRNA	complement(join(2844..3521,3590..4619,4684..5341,5819..5866,6647..8156,8220..8346,12506..12893))		
	/gene="BcDNA:GH06348"		
	/locus_tag="CG1516"		
	/product="CG1516-R"		
	/db_xref="FLYBASE:FBgn0027580"		
mRNA	complement(join(2844..3521,3590..4619,4684..5341,5819..5866,6647..8156,8220..8346,12506..12893))		
	/gene="BcDNA:GH06348"		
	/locus_tag="CG1516"		
	/product="CG1516-R"		
	/db_xref="FLYBASE:FBgn0027580"		
mRNA	complement(join(2844..3521,3590..4619,4684..5341,5819..5866,6647..8156,8220..8346,12506..12893))		
	/gene="BcDNA:GH06348"		
	/locus_tag="CG1516"		
	/product="CG1516-R"		
	/db_xref="FLYBASE:FBgn0027580"		
mRNA	complement(join(2844..3521,3590..4619,4684..5341,5819..5866,6647..8156,8220..8346,12506..12893))		
	/gene="BcDNA:GH06348"		
	/locus_tag="CG1516"		
	/product="CG1516-R"		
	/db_xref="FLYBASE:FBgn0027580"		
mRNA	complement(join(2844..3521,3590..4619,4684..5341,5819..5866,6647..8156,8220..8346,12506..12893))		
	/gene="BcDNA:GH06348"		
	/locus_tag="CG1516"		
	/product="CG1516-R"		
	/db_xref="FLYBASE:FBgn0027580"		
mRNA	complement(join(2844..3521,3590..4619,4684..5341,5819..5866,6647..8156,8220..8346,12506..12893))		
	/gene="BcDNA:GH06348"		
	/locus_tag="CG1516"		
	/product="CG1516-R"		
	/db_xref="FLYBASE:FBgn0027580"		
mRNA	complement(join(2844..3521,3590..4619,4684..5341,5819..5866,6647..8156,8220..8346,12506..12893))		
	/gene="BcDNA:GH06348"		
	/locus_tag="CG1516"		
	/product="CG1516-R"		
	/db_xref="FLYBASE:FBgn0027580"		
mRNA	complement(join(2844..3521,3590..4619,4684..5341,5819..5866,6647..8156,8220..8346,12506..12893))		
	/gene="BcDNA:GH06348"		
	/locus_tag="CG1516"		
	/product="CG1516-R"		
	/db_xref="FLYBASE:FBgn0027580"		
mRNA	complement(join(2844..3521,3590..4619,4684..5341,5819..5866,6647..8156,8220..8346,12506..12893))		
	/gene="BcDNA:GH06348"		
	/locus_tag="CG1516"		
	/product="CG1516-R"		
	/db_xref="FLYBASE:FBgn0027580"		
mRNA	complement(join(2844..3521,3590..4619,4684..5341,5819..5866,6647..8156,8220..8346,12506..12893))		
	/gene="BcDNA:GH06348"		
	/locus_tag="CG1516"		
	/product="CG1516-R"		
	/db_xref="FLYBASE:FBgn0027580"		
mRNA	complement(join(2844..3521,3590..4619,4684..5341,5819..5866,6647..8156,8220..8346,12506..12893))		
	/gene="BcDNA:GH06348"		
	/locus_tag="CG1516"		
	/product="CG1516-R"		
	/db_xref="FLYBASE:FBgn0027580"		
mRNA	complement(join(2844..3521,3590..4619,4684..5341,5819..5866,6647..8156,8220..8346,12506..12893))		
	/gene="BcDNA:GH06348"		
	/locus_tag="CG1516"		
	/product="CG1516-R"		
	/db_xref="FLYBASE:FBgn0027580"		
mRNA	complement(join(2844..3521,3590..4619,4684..5341,5819..5866,6647..8156,8220..8346,12506..12893))		
	/gene="BcDNA:GH06348"		
	/locus_tag="CG1516"		
	/product="CG1516-R"		
	/db_xref="FLYBASE:FBgn0027580"		
mRNA	complement(join(2844..3521,3590..4619,4684..5341,5819..5866,6647..8156,8220..8346,12506..12893))		
	/gene="BcDNA:GH06348"		
	/locus_tag="CG1516"		
	/product="CG1516-R"		
	/db_xref="FLYBASE:FBgn0027580"		
mRNA	complement(join(2844..3521,3590..4619,4684..5341,5819..5866,6647..8156,8220..8346,12506..12893))		
	/gene="BcDNA:GH06348"		



Db	236028	TCGGGAGGGAACAACCAAGTACTTTGGCATCTTCAAGTG	236071	
RESULT 12				
PFMAL3P2				
LOCUS	PFMAL3P2	158548 bp	DNA	linear INV 29-JAN-2003
DEFINITION	Plasmodium falciparum MAL3P2, complete sequence.			
ACCESSION	AL034558	AL008982	AL009007	AL009008 AL009009 AL009014
	AL010140	AL010150	AL010152	AL010156 AL010158 AL010211 AL022217
	AL844502	297349	297350	298546
VERSION	AL034558.3	GI:15375364		
KEYWORDS	HTG.			
SOURCE	Plasmodium falciparum 3D7			
ORGANISM	Plasmodium falciparum 3D7			
REFERENCE	Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.			
AUTHORS	1 (bases 1 to 158548) Bowman, S., Lawson, D., Basham, D., Brown, D., Chillingworth, T., Churcher, C.M., Craig, A., Davies, R.M., Devlin, K., Feltwell, T., Gentles, S., Gwilliam, R., Hamlin, N., Harris, B., Holroyd, S., Hornsby, T., Horrocks, P., Jagels, K., Jassal, B., Kyes, S., McLean, J., Moulé, S., Mungall, K., Murphy, L., Oliver, K., Quail, M.A., Rajandream, M.-A., Rutter, S., Skelton, J., Squares, R., Squares, S., Sulston, J.E., Whitehead, S., Woodward, J.R., Newbold, C. and Barrell, B.G.			
TITLE	The complete nucleotide sequence of chromosome 3 of Plasmodium falciparum			
JOURNAL	Nature 400 (6744), 532-538 (1999)			
MEDLINE	99376085			
PUBMED	10448855			
REFERENCE	2			
AUTHORS	Hall, N., Pain, A., Berriman, M., Churcher, C., Harris, B., Harris, D., Mungall, K., Bowman, S., Atkin, R., Baker, S., Barron, A., Brooks, K., Buckee, C.O., Burrows, C., Cherevach, I., Chillingworth, C., Chillingworth, T., Christodoulou, Z., Clark, L., Clark, R., Corton, C., Cronin, A., Davies, R., Davis, P., Dear, P., Dearden, F., Doggett, J., Feltwell, T., Goble, A., Goodhead, I., Gwilliam, R., Hamlin, N., Hance, Z., Harper, D., Hauser, H., Hornsby, T., Holroyd, S., Horrocks, P., Humphray, S., Jagels, K., James, K.D., Johnson, D., Kerhornou, A., Knights, A., Konfortov, B., Kyes, S., Larke, N., Lawson, D., Lennard, N., Line, A., Maddison, M., McLean, J., Mooney, P., Moulé, S., Murphy, L., Oliver, K., Ormond, D., Price, C., Quail, M.A., Rabinowitsch, E., Rajandream, M.A., Rutter, S., Rutherford, K.M., Sanders, M., Simmonds, M., Seeger, K., Sharp, S., Smith, R., Squares, R., Squares, S., Stevens, K., Taylor, K., Tivey, A., Unwin, L., Whitehead, S., Woodward, J., Sulston, J.E., Craig, A., Newbold, C. and Barrell, B.G.			
TITLE	Sequence of Plasmodium falciparum chromosomes 1, 3-9 and 13			
JOURNAL	Nature 419 (6906), 527-531 (2002)			
MEDLINE	22255708			
PUBMED	12368867			
REFERENCE	3 (bases 1 to 158548)			
AUTHORS	Lawson, D., Bowman, S. and Barrell, B.			
TITLE	Direct Submission			
JOURNAL	Submitted (17-DEC-1998) P.falciparum Genome Sequencing Consortium, The Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, UKXX			
COMMENT	On or before Aug 30, 2001 this sequence version replaced gi:2982545, gi:2982529, gi:2982547, gi:2982546, gi:2982548, gi:2624372, gi:2982553, gi:2982560, gi:2894487, gi:2665315, gi:2982565, gi:2982575, gi:2982578, gi:2982517, gi:2982518, gi:2982519, gi:4493878. For more information about this sequence or the Malaria Project, see <a href="http://www.sanger.ac.uk/Projects/P_falciparum">http://www.sanger.ac.uk/Projects/P_falciparum</a> .			
FEATURES	Location/Qualifiers			
source	1..158548 /organism="Plasmodium falciparum 3D7" /mol_type="genomic DNA" /isolate="3D7" /db_xref="taxon:36329" /chromosome="3" /clone="MAL3P2"			
gene	join(1030..2337,2429..2502,2645..2685,2826..2941, 3046..3173,3308..3388,3508..3571,3699..3737) /gene="MAL3P2.1"			
note="synonym: PFC0160w"	join(1030..2337,2429..2502,2645..2685,2826..2941, 3046..3173,3308..3388,3508..3571,3699..3737) /gene="MAL3P2.1"			
note="Revised: revised to 8 exon structure, revised: donor for exon 7 4 weak Pfam matches to entry PF00023 ank, Ank repeat; Pfam match to entry PF01529 zf-DHHC, DHHC zinc finger domain, score 66.70, E-value 4.8e-16 Similar to Plasmodium vivax pvhlh4090_p pvhlh4090W SWALL:AAF99464 (EMBL:AY003872) (611 aa) fasta scores: E(): 5.6e-172, 71.94% id in 613 aa, and to Mus musculus 5033406114rik protein 5033406114rik SWALL:Q9D308 (EMBL:AK017158) (368 aa) fasta scores: E(): 1.7e-09, 24.87% id in 197 aa, and to Rattus norvegicus small rec srec SWALL:Q9UKR5 (EMBL:AF228917) (366 aa) fasta scores: E(): 8.4e-09, 27.95% id in 161 aa"				
/codon_start=1	/product="binding protein, putative"			
/protein_id="CAB38988.3"	/db_xref="GI:23477021"			
/db_xref="SPTREMBL:O97224"	/translation="MNDNESLDSLEVKQYEIIKAKYQDFIRLQILIQVLLNNDIE MNSINILHWACYCGTFELVKLISFENEDKLEALHLELLYMNVSLEQNEHGQTA KCFQISILPHKRRRSPFLTAICEFNEDKLEALHLELLYMNVSLEQNEHGQTA LFLSVKNNISTLQWLLTKENVINRDFYGNVTLVHIAVRHCDIDILRLCCYGCGLNV YSSIEKNNTNVFQLCIKRNVFLVYILLKWLQNKIKSLKICKTIYAFVFWFPAIL NLIVYFNIAHSPSILINKYHFKSLIWTIWFQQFLWCLYFKSPGFYKHNHMFKNKK NIOSSKTFNSPMWINGTFKNAEYQNLNIEEIFQINKKLISNFNLINPHDQALS NKYNDILINLEYQLSLYSQVSEIRNSLSDYRNALYNQNPNCIVCTNIIIPRV HTCAECPHCIVHQDHHCWVDCIGIKQRCFYMFICFVLLNYTYVYLYPHLFO AIIYAFGSLVILCNFVTLFAFITYKLVFARNTILTNTVFVHYHKKPNHITDKYNT DACWDQFQNLNLGILKNIFYFWSLNVDFYIRQCTKTSNDVCYTLISNT"			
5341..12021	/gene="MAL3P2.2"			
/note="synonym: PFC0165w"	5341..12021			
/gene="MAL3P2.2"	/note="Revised: original entry split into PFC0165w and PFC0166w"			
Similar to Plasmodium vivax pvhlh4095_p pvhlh4095W SWALL:Q96215 (EMBL:AY003872) blast scores: E(): 0.0, score: 2249.43% id, and to Plasmodium falciparum thioredoxin-like redox-active protein fred SWALL:Q9NC62 (EMBL:AF234633) (179 aa) fasta scores: E(): 7.5e-47, 99.42% id in 175 aa"				
/codon_start=1	/product="hypothetical protein, conserved"			
/protein_id="CAB38989.3"	/db_xref="GI:23477022"			
/db_xref="SPTREMBL:O97225"	/translation="MNRREKLILKEKQKRNKNIKNVFRDCHCMFNYSQCIVNVK NKYENDFWKYEINIGNDNIKSVQVEEKKKWIIDNYNKRKSELETPYRYESNRN IFIILINSFENLNGSCVYNNDLIYIHNKRVEHMLNPKRTKKKKKKKRTKTRK KRTLHNKNDTFGSVTSMNDEKQYKQYHDKDDDDDDDDDDDDDDDDDDDD YEDDDDDDDDDNNYNSYDQYNTKHNKKIKINPMYNSITPFSNNFVCKDB KRKKEKQKNSIDKRNIIYSDSDNNYEHIFTHNSDLFSLSDHIEKNKLLYQ NEQFNYDDVIFRKFHKNLFWNMSCNSWIVIKSILINEYITRSQKQDLVDVND LIDVVMYHRIYNYCTRLYINDDFLLNNKTKELGVQNVASHINNTNNMNDNI HNVNINNNMREHLKIEKKKNNNNNNNNNDDELSTTDTGDDMLSNCCYNEE SKYKLVLIKMTLKHLFNQTDINILFSFNACLSYKNKLDENSYLLINNYFNKK SEHLINTSRKNKVTLDHTNTYNSNTNIIKTYNSNNNNNNNDNLISNDLYSYFN VLTKIVENYSTSLYIKECLLIDILYILNNTYGNLIYIKYKYDHYENYTNII IVQKKNEKKKKSKYKNIYHSHKNNRQKNSIYSSFSQKQYKNTNMNNNLHGE EIPDSSSSNNSSNNSSNNSSNNSSATSYDIPKMRNNKNNLYINRQKNDLLY FPFKYLIHIPTLNNQNNNSIETAREILYGIYIRKIQKIEKNRQSCNFTFD IFCCCLNSISAHILNHNRIEENVRNIYNTKINMKHNSNTNNNNNNNNNNFSLY HNDDLCLDFEYKLSCTDPEEKIIPFSINLPHKNIQDFNLFPFNIEKAEQISLY KIKISVLPILQIGKLANKINKIILKKKINNSKYLIDILYVLOEYINPDDINKVLY LKNITLTFDFIKNYIFYGKQVQDYKEFIHENKHIPTPYKQKHKFPNNMYKIFK KYIDLSTNYWGAKYIILNDKVPFLKLSNHFITGKYIDVLKSACKILNIQRKTSFY LIDQNEYKNINENIDNMDTNQNNRNSFLPKQVYIKRIYENINNDQYNSDHRNP YFYNDLYKSERQTHQGMNKIINGKVKYNTMNLPEYDEYEHKNNKDEKNSCFLIYDN			





gene	/chromosome="5" complement(219..8039) /gene="PFE0420c" complement(219..8039) /gene="PFE0420c" /note="Similar to Plasmodium falciparum guanidine nucleotide exchange factor pfirc1 (1327 aa) fasta scores: E(): 9.6e-190, 97.45% id in 1338 aa" /codon_start=1 /product="guanidine nucleotide exchange factor, putative" /protein_id="CAD51450.1" /db_xref="GI:23504571" /db_xref="SPTREMBL:Q81418" /translation="MYEODDEDKDIHKYIFRPFPKISTGLDHYACVAYKKNPSIFTW GSGNSQIQLGINRFPVNNPTVCFDFSPFIVASICTNYATFVLVKVNDIGCSVYS FGKGNGLGKKKRELNFLEHEEKNIQVKEQLOKQVNDPLNQLVDAFGIKKKE KEKKKKKNDI1KKEKEENNIINNEKNNNNNNINSEKKKNDINGEAKYVQNK KEFYEHNFKNDIGLPSDDDIYDDNTSGDESLLNFADNIISSIQNDETWMEDVYNN IVNDHTNDRQSDINETIDKSHDNDYDDEKKKEMNIYKNEKLYSNIKENDNKHDW FTPPFMKIVFPKNTTKIFISGDMHTLAICTGGVLSWGVNFGCIGNTNQNVYEP TPIFLEHSDVNGNDNDYDDNNNNSTYKPHKEDSLKWDKKKNVTHCSAGSKH SLACNLHGDIYSWGVGNGRGLGNIKSYNKPQLIKGLNRKRIYVACAGSHSGCIDS NINYTWNGGFIYKLGNDNDIYYPKLVCELNKKVCMWSEFCNSALNAGKGV YWGTFNTNNYVYISKLPKQINTNYKCISSIAHSTYVCFGLTVGLDLSFGYKNY SYKEINDNDSDDLIKYFINKEVEEQHYNGDNKSVHTKNNIISSIVKSKDHEFDI HYIKEMRGKMYIKDITVHFYNDTWYTMNKIKHIDNVTLNNDNYVTYDNTNGSHG NIFEKNLKSKVKI1DGSDFRSVELLESKVYTAGYKNGELGNGEFLNKKRGVP ILLNICYNKII1KIAGGNYVLCVSDGLVYWGKNDKSQIGIVIKDFYEPVHVKNLT NVNIFAGYDHSACILNKLKDKNTGVNNVNVNGVNVNGVNTSEDCGDLIYWGNAE SGKVLGVYTGCGCILLPDKNTINKTKYCSLNGSHSLFTNTNELYVCGSNNGRLG LFEKSKVCNTYVNTVYKNNIYIKDILAGNTYSIILSDVGFYIIFGFIKSGNRL SNMKSCEISYINTEVTKVMNGKYEHVFTLDYDKLFGIKNKNCQI1LDCNKENNY IKPKPLIYFLKNNKIIESTFSPHNVTFIQFNNSIETPANGYTNVTHLCIGIPNNIKY LKHPKIIKTLWTYDQKDLVNGSCDMELNENYNTNMNYNYIKYINIEITLVKNI NTTELQNEYTYFGYEEIEENFIYINVINNNIYNWCHLQNLKKEEYNNMYI1KSP EKDILIDYKHEIYVLLNIYEQFNLYNLQFNLSTANMKEQHLNIPLYPFTTY IFDQRI1KLEFYII1QQPFLYI1LCLIHNYKNVKNQMLTKGMSYNNNTTHIY HNNTEQH1KNNYNIINSNYINHTYEQNDALKYNSDI1QNKYSFYKNSNVLC FIDLYADFRNERVTI1FTI1KLAGIEEIKSNLNSH1ENIDTS1FPTI1OMLFPFK EILLNFAOCLLDLNPNSFVTL1DLSRKVPKNI1PLNQGVSNAMQHTKLONT PMYGAHRENL1LNDYNDVNDKNIHVNSHTMDLNGANKKAKK1KKNKNNKK NNKKNNKNNKNNNNNGGDDYSGDDIYSGDDIYSGDDYSGDDYSGDDYSGDDY SGDDIYSGDDIYSGDDIYSGDDIYSGDGINSHVKNLKNKNVREPNISDFIDNLD EKDFEFLNFTKGQTSNEDKRLPVNDISCVKSGVPYNI1PINNEDENMYNRSQEI QENLNDNDNVCI1NFCV1PEELCKI1PLNI1KFPDMFKI1IKELFKYFI1YEKNLNDI1 NOTWKIYVQNHHTIYI1PFPKLI1LMAI1N1LKNMENIAHKSYPNI1FPHILNIRKI CDFEILYLKYESLN1YNI1S1KII1FHTFISLTN1NNICNTHHHIYDAYTNL FHYHLNRNSIYVQLK1FQVCH1ENLFFRFQNYLMLSFNDPALD1VNFYFYKNIPIQ KHGHKHTLNSWTGNTNMDI1QKERH1NL1KGLGDMGNT1QRPKGVSTFTGHQDNKI NSNEGNSLLARKAKNDLND1R1FEMKEGK1NG1VGMKRGDNPMPGYNTPMGNNPMGGYN PMGNNPMGGYNPMGNNTPMGNNPMGK1NG1VGMKRGDNPMPGYNTPMGNNPMGGYN AFNL1PSTIKITSNDNKKRKKMKRL1FSEDDI1EPFKCL1YVVKLDIR1ELLEKKNMSI CEFTK1PMPQYICVRKQTL1ENNEYL1FSI1HKYNEKNNIYI1SECLKNCPI1LEHCTD TNNLILKLMKMYMSLDQKQNNL1LH1LKAID1FLSDEMTYVEFLDFEPANLSIY KFKDDKTLKSNMEQYIFDMLONTYKNSFFQVKNRN1Q1SLN1LRKKKHMNYLKK LYEKQYIEIDL1NIYQHKMKD1VTLKNAI1FVSKLYIEKPI1LLSHSYKFNKLNFVFL KKKHQIORTKFLPVNSSTVH1YQMKDILNNOIL1NTH1ETLYTLVWYLT1E1FEDLNNI IKFSLI1LTDQNNKI1NEHT1FTGSEIYSVMSN1SPYI1YFPFKYSKYLCSITGLHPMH ILLNFVTDLY" 9222..10010 /gene="PFE0425w" 9222..10010 /gene="PFE0425w" /note="Weakly similar to Arabidopsis thaliana F8n16.20 protein f8n16.20 TR:Q9V2V3 (EMBL:AC005727) (332 aa) fasta scores: E(): 2e-12, 28.763% id in 299 aa" /codon_start=1 /product="hypothetical protein" /protein_id="CAD51451.1" /db_xref="GI:23504572" /db_xref="SPTREMBL:Q81417" /translation="MFSKSGRVKRPADNRLPVASLKTSEIWHKNSVGYPDYASYEIN KKKEDKDEDINEKAKNLFNLSRUTGITSTTIPGACTVCNHHGLPQCRNF1SLEKFD
CDS	
gene	INNMNNBEDDIKERKNILGLMSDDGSDSHSDSHSDSHSDSDSDSDSDSDSDSDNSNS HSDNSNSHSDYSYKKKKSHHSHSKYKKKSHHSHSKKSHHSHSKKSHHSHSKKSHHSHNDTHK SDEKHKYISSKAKKIKKBEKKKKYHSHKMKKNNKNNKNNKNNKNNKNNSSD" 11261..15733 /gene="PFE0430w" 11261..15733 /note="PFE0430w" /note="Similar to Rattus norvegicus RNA helicase TR:Q62780 (EMBL:U05746) (1032 aa) fasta scores: E(): 2.4e-59, 38.387% id in 719 aa" /codon_start=1 /product="ATP-dependent RNA helicase, putative" /protein_id="CAD51452.1" /db_xref="GI:23504573" /db_xref="GOA:Q81416" /db_xref="SPTREMBL:Q81416" /translation="MYNIFNYIKDSKKEKKNDKDDMIKVKKQDHSDSDVLRKDN ISNKKKQDVERNYTVDSYNEEKNKKKKKKRDKSKYSVDDSYNKTQNIYNDHKTDE YPSYAKKSSYHHVREKYPDKRRKYEH1GKNKRSYDYSSNYDESDFSSSY YNERKKYKYEDKKKEKEKEKEKRRKRYHHYDDDDHLEYSYSSRSRSGSIN IKBKQHKKLYSHEKRRKKKYSNSAYSSDEETKKEKKKKKSHHDFVLDKKNK YNYNDEETNKQDENNNNNNNNNNNKDSSENLNKTHEQREKSNINNKISKEN QBI1CEENKEI1QKVENNSNVKVI1TENNLSRLERLKL1FARKIKTNDTNKTKFT LNTVKNIKNLNIFINDQDESNGKEFDI1KNEKNINQVNDVNNKDDTKK1YE NHKIKEDSNTNDIHVE1NNKNEDAL1FMKEI1BOACEEKKNVKKSL1DLEI1YTD MKNQKYPEINPKNEE1KEYNAKYGNII1QGN1HEKQDSNVNTQEDLKKNKNTK NTNNYSDNNNNDDDDDDDDVDFHKL1FVEELKKTEEDKNEKNOMNESLTD SNKKNGKGSDFSSDDTD1DKLSNSELN1EENGKMMKNKLEVNHDHDIYPI KNFYVQKEI1TMKDSVDMPFRKNGNI1VKGKCPRPQYFQCGLPK1LQILEK KFKMYNI1QMTI1PALMCGRDI1A1AETSGKTL1YLPV1R1VH1HQB1PLRNDGPI SI1LTPRELS1QVKNKAI1YKAVN1E1LA1VYGSGNIARQ1KVL1KGV1L1VGPGR I1DLTI1SNCKVTN1NRYSVFVLD1EADRL1DL1GSG1QYLN1L1RNC1KQ1TAM1SATP PNYIQMAKKLY1KPIE1I1VGEK1KTNNI1QFVEI1EESK1FRL1LGL1GWI1KYL VL1FVQK1EADLY1LELY1DYND1LL1VHGQD1D1QFT1EKFKEENK1VLI1AT1VM ARGDII1N1I1NY1QCPD1H1EDY1TH1R1GTR1SN1NY1AY1F1L1P1NY1KAY1D1Y1NL LKNY1Y1AKT1DI1D1Q1DEN1M1E1Y1T1K1S1NEK1QK1N1G1K1G1K1Y1P1D1E1K1SR H01DALAKELG1LQ1ENDENK1LOV1DEL1GNI1EHNNT1V1CN1KN4DN1K1YD1E1 DKESNI1PY1T1HNSNNMFEDK1SGI1SKNSKNN1TNEK1KNT1QY1D1I1TKV1BO1I1Q I1KNDT1MEL1SKAKK1NELM1KTYN1FLS1QENH1DK1LNA1Q1E1F1DK1IAE1K1AL1K1VTEH I1KDEK1KLI1FNK1TED1IKK1LI1QNI1QSD1TNES1QRN1MLLS1ALUR1KNMKK1SYLP1 HTY1TD1Q1LH1E1I1APPY1NQHVRLK1SHK1VLAK1ADM1SGAT1C1Q1G1Q1SN1Q1PNK QNF1LD1Q1LH1E1I1APPY1NQHVRLK1SHK1VLAK1ADM1SGAT1C1Q1G1Q1SN1Q1PNK complement(16598..17442) /gene="PFE0435c" complement(16598..17442) /gene="PFE0435c" /note="Similar to Rickettsia prowazekii single-strand binding protein ssb or rp836 SW:SSB_RICPR (Q9ZCC2) (152 aa) fasta scores: E(): 1.1e-16, 38.15% id in 152 aa Signal peptide predicted by SignalP 2.0 HMM (Signal peptide probability 0.896, signal anchor probability 0.017) with cleavage site probability 0.471 between residues 25 and 26" /codon_start=1 /product="single-strand binding protein, putative" /protein_id="CAD51453.1" /db_xref="GI:23504574" /db_xref="SPTREMBL:Q81415" /translation="MGKMLCIVFPLLIYFNVLIHRTYGYIGDVVKVHQLNNI1NKR1 IKSRK1SMFKNLQNDNFYENKFRYNNNNNNMKNLIMLI1GRVCEPDI1KLNG GDKVATFSLATNEFRDNTNELKSTDKWHI1VYVDQNI1VDLI1DKYLRKGRVYVQGS LHTRKWH1NDMNSQPKQI1E1LSYNGKGLI1FLDDK1RNFNNNNNNINSNQOH1NN EH1NNNN1NNGND1FMP1LNSDKI1EDKEFT1DLDDNNNEENN1FQSNSET1FDKQEGI1YD MVVQEFEE" 20022..29426 /gene="PFE0440w" 20022..29426 /gene="PFE0440w" /note="Weak similarity to Plasmodium falciparum Pfc0425w protein pfc0425w TR:O77336 (EMBL:Z98547) (4550 aa) fasta scores: E(): 7.2e-23, 22.792% id in 3238 aa" /codon_start=1 /product="hypothetical protein"
CDS	



1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
84

Residues, Lack of Direct Relation between Consensus Residues and Transmembrane Helices, Expression Patterns of the Transport Protein Genes, and Protein-Protein Interactions with Other Proteins.

5 (bases 1 to 69009)  
Sethuraman,A. and Dolinski,K.J.

Direct Submission  
Submitted (20-DEC-2002) Department of Genetics, Stanford University  
School of Medicine, Saccharomyces Genome Database, Stanford, CA  
94305-5120, USA

Deletion of 16 nt (GATCTCTCTGATTTT) between 1437732 - 1437747  
Sequence update by submitter  
6 (bases 1 to 69009)  
Sethuraman,A. and Cherry,J.M.

Direct Submission  
Submitted (14-OCT-2003) Department of Genetics, Stanford University  
School of Medicine, Saccharomyces Genome Database, Stanford, CA  
94305-5120, USA  
Sequence update by submitter  
On Oct 14, 2003 this sequence version replaced gi:27316972.

Sequenced by:  
Stanford DNA Sequence & Technology Center  
855 California Avenue  
Palo Alto, CA 94304, USA

Curated by:  
Saccharomyces Genome Database  
URL: <http://genome-www.stanford.edu/>  
e-mail: [yeast-curator@genome.stanford.edu](mailto:yeast-curator@genome.stanford.edu)

Neighboring Sequence:  
The 5' end of this sequence contains a 1000 bp overlap with GenBank  
Accession Number U33007.

Location/Qualifiers  
1..69009  
/organism="Saccharomyces cerevisiae"  
/mol\_type="genomic DNA"  
/strain="S288C; AB972"  
/db\_xref="taxon:4932"  
/chromosome="IV"  
/map="48"  
1..12144  
/organism="Saccharomyces cerevisiae"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:4932"  
/clones="cosmid 9410"  
2618..38763  
/organism="Saccharomyces cerevisiae"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:4932"  
/clones="cosmid 8035"  
38274..69009  
/organism="Saccharomyces cerevisiae"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:4932"  
/clones="cosmid 8166"  
68472..69009  
/organism="Saccharomyces cerevisiae"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:4932"  
/clones="cosmid 9787"  
521..10327  
/gene="TOM1"  
521..10327  
/gene="TOM1"

/note="YDR457W; Hect-domain-containing nprotein, required for G2/M transition; similar to Rsp5p, ORF YER125w, from SwissProt Accession Number P39940, contains motifs typical of protein kinases; CAI: 0.16"  
/codon\_start=1  
/evidence=not experimental  
/product="Tom1p"  
/protein\_id="AAB64910.1"  
/db\_xref="GI:9277738"

gene  
CDS  
/translation="MVLTRCEKARKEKLAAGKPLVDVLIIDCDTPTFLERIEAIOEW  
DSRDDLYWPIIDLRMDGLLKVARYKYKQPKKCEKVELEMAHVDVYCLKMLK  
FTRLLNLTENRFPVSSGDMVLLNCPNFTIKLAVMRILAILGKRVFVAREKIVAHN  
IFGDHNRKTKLALSLSSVMDEGEHFLVDYFKKKYQPKWRKLFTHYTSND  
FKKSSQCKNNINETSITKVTMTTOELCEHSLQOIFDKGALLPAEDFYSISAKV  
AKAFSDSGENIDLRNIIKETKLNATFVNTIIFSPQVSSKPELDPYAFNSITDLIS  
LSETKIPKELRTDALFTLRCISLKHVWCSDIIRNLGGNISHGLLQILRYIAKTLREA  
THEIDREYNVRFYLIISNLADYKPLHESLFAAGLIPTLLEIVSINCPYKRTIASATH  
LLETIDNSETTEFEIDGFTMLITSVANEIDFTLAHPETWOPPKYSVYVYSISFRE  
LAYIRSLKLVLKLSLTDGSDIRNLIDSPILVSLKLLLENKLFGLTILTYTLDVVQ  
KVINSEPTIYVLEAGLIIPYVIDNPKLIGLSAELLSPDWVAIAICNPLGKQVK  
EKGLINLFDLLADHARILGTGDRSTEYQIDELARHYDOLKANIVEALCNVTRK  
MSTPRNREFELTSPKQOKYFHRKNEILTDKEHEPAYWELDKGMPLTFTSVL  
FGMSLNGSFQVPOHLEARDPLAIIIPENPPEYFTSVAISNTEVLYLDEKEDY  
AFDMVKLVNDQENLNDLNFSPDRSFLEDRGNSVNSCHSKLRALAILNVTNV  
YIDLTTLSCKRIMQIYSYFDRKGFSLIKNLKLLQKCALCEEMYIRQHPDVSITMP  
LPIVDVSGDPLQIYIDDKKGDQKIKTSVTRNTLQMTLITLQASNTAIFRCF  
LRLSHRNMDLHKLTTEVHFENNVENVIMLKATELGHLPIVLNLNNTVFVT  
IPKAPNSTEILQITPAYIFYOKGGVLLVHIIRDLFTMTKI KOLSLDNIYIDES  
NGILTSLCLINALTFVNSMOTEMVQSIKYVYISDIDDDYINMKALTVPYIKWALA  
MILDLKXSLFKTSRNVYISVFKQLLSMLKNIITINVINIYKELTEHMDLIIFPIK  
KISLFEQVIGPGDAANYLITDGDLPADNSIGLFSBPQWEKYKKLIOEDKSIYYPQ  
MQAQYKGCSEKDELDRDTFENDGLPSRIFTLPFKLVNAFAKTLIQIFTKYDEP  
TEVFAGRILDRILETDLPATLSLILHFGILFNEKYIYOKASHMORFIYLBKSL  
KPEHVNTPMFSKALYYEILAKSELPHLELSKDVLLRYPLLSNAKVFRIIDPMKOK  
LFDILIRVSDISNFYSALATSRILIFYSRDELYANNIARSGILSKLVKIGSFQKLDK  
INFLSSFLLLTRCFETTFENVDALIRAEINKSFARPLGGDKDVAURELTILSEKAF  
VVMRSPQFIDVLCETARPEHFDQDALVDYSLKRFGLGKDKNTQASSTESKDIYERT  
GIMHLLLSQMAASEKMDLSEPAKSDILPENKKAQLDRSPNVCAYMIFLLKLLVELV  
SSYNCKFEFTFSRNTYAERPRRTAINFLYRLLDKPVGTDDHKEAKRRVIG  
MLARSVIIGLATVQDDRTTKDVKLADPHMNFIRKFAEAIKAIKRNATSSKILLES  
NHLKLDMPFRIITSMVYQAPYLRQLLDSNKVEADQYQCLKVIDLGLPSVITEMAS  
IDLNTPFKSFNFVAEALNTISRNPFSEHFKIDHDEVEDEVEDSKDEIPDMFK  
IDNLGMYDVEDIEEDDDDTSLIGDDAMAFVDSNGFEVFEVDEDDMDGEDADAR  
SDSENELSSEMQSTADGTDVDEYDADGLINIDQSGDEDEMDADANISSHSS  
SENEDDASMDVIEVYDSDYDLDSDYDVEDSDMDSGLSLSISDESDSSDEP  
INSTRMGDSRRRLIAEGVELTDDSQGESEEDRQVFRGIHIFSENEPLFRVHDEM  
RHRNHRISINRTHFSAMASPSLSLLNRRNQSLINPLQPTGLEOVENDISDOVTY  
AGGSRPSRSHLHFSEVLVSGSPFDEPVLGILLKSTVSRWKIDFMDYDSKTYANCI  
IPTVINRLYKVSALQKOLENKEQEKLNKLLNLFNEAKVSHNSSDAISVEQDDIQE  
SNVTHDHEPVVTIIGSEVVDIGDIDPFPMNALPDDIRADVFAQHVRRERAEARLN  
SDNHVHSREIDSDLEAIPEDIREGILDEAEQMFGRIGSADVIRADDDVSNDE  
EVENGLDGNNDNRNADPEKKAPARIYFAPLIDRAGIASLMKSVFISKPYIQRIYH  
ELVFLCSCQKQNRNDLMTLFIKLEGIIDQHSLEKYNIISSRAMGHAKTTVRQLP  
SDCTPLTYANOTIEILQSLIDADSRKYFLAEHNLIVNKANNKSRKEALPDKKLW  
PLWHLFSLDKRLITDESVMDDTLRIQVCTKLAVLSTSGNKENSLKXFLPSFDF  
EDDLMLKLSIIMLSDCTTRVFQOTLIIYNSLKGQCMSIFTKHLVSLAISIMSKLS  
ALDGLSREVGITITGMEINSELLQFTLPSDQAKLLKLTITVDPLYTHKREBERNV  
KDLQSLYDMNGGPPVMSLSLCSLQFEKSOAINTSATILPLIESLMVVCRRSDLSN  
RNTAVKRDALKLLDKSTRVENLPFTDAHKLLNOMIRSNPLKMSGPEALLVKNPK  
VLDPNKBYPFNALKSDNQSRPKLPITVREOVPLDSYRALFPKTNDEIKNSKLEIT  
FKGESVDAGVTRREYQVLSROMPNPYALFLVPFSKTIPTFHNRTSGINPEHLSFF  
KPIGMIIKAIARDQCFLDCHFSREYKILRPSVKDMESLDPPYKSLVIMLENDI  
TDIIEETFSVETDYGHEKVINLIEGGKDIIVTEANKQDYKVVYKQLTSVKRQMD  
NFLVGFYALIISKDLITIFDEQELLELISCLPIDVDWKNNTTYNYTATCKTSYFW  
RAVRSFDAERAKLLQFVTGTSKVLPGFKELSGVNGVCKFSIHRDPCSSERLPSHT  
CFNQLNLPVSYETLRGSLIILANEGHEGFLA"  
Complement (10786..12777)  
/gene="YDR458C"  
complement (10786..12777)  
/gene="YDR458C"

/note="similar to Saccharomyces cerevisiae hypothetical  
protein on chromosome XIII from GenBank Accession Number  
S49745, CDS number 5; CAI: 0.13"  
/codon\_start=1  
/evidence=not experimental  
/product="Ydr458cp"  
/protein\_id="AAB64934.1"  
/db\_xref="GI:9277762"  
/translation="MDHRLDPKTLKVSQRLVLENDVAFANARKPVLKLPKPK  
QRLQSSPEASKVTSIQVVKSAKADRKTKLKSKKLESSSESSTKVKDENVETNK  
RRREQIHTDNEAKQIOEIKSPKKRKRSSKANKPPSPQSGKATSDLTSEL  
ETVELHKKSDDKPRVKELPPELNLKVSNEFLAQLNKLASAAATENDHGIKST



DLSSRIETEPVGPSTGAETRNSESEVENINLEVPQEVKEBELTKISTFPDQDE  
 EDTSLSSKNSRIPKGRHFIAKTKRGIDIMKPFIAHLFIWLNAGAFSLGICPI  
 LFGWYRQRIOVGCHEKPLKSLATSAFFQTERVDSVLQAYRPNCLCEPHEGICSS  
 FMNCEPGEYKPSILTYGIIPPKYCAKDESKEDVLELVKVNVEYLKKNAQHE  
 CGEENIPESGETETKLYDIFSHSRPWSHESORENDKWNVLEILKKDDIILWLPDR  
 ETNGKREKSNNTIYIRTSKRWVLCQLEGGDIQEIYITIKIGSFLITUGVUFLIK  
 KIQSTLDNYVQGEIIEKLKVEADIKLQKVNKGESFPFLTVQLRATLLSDIPNIKE  
 QNNLWAOPTKIMKEQSENIELYLLEENGEIMTCWEKWE"  
 complement (13050. .14174)  
 /gene="YDR459C"  
 CDS  
 complement (13050. .14174)  
 /gene="YDR459C"  
 /note="similar to Saccharomyces cerevisiae hypothetical  
 protein from PIR Accession Number S51289; CAI: 0.12"  
 /codon\_start=1

## Alignment Scores:

Pred. No.:	0.841	Length:	69009
Score:	165.50	Matches:	78
Percent Similarity:	35.73%	Conservative:	66
Best Local Similarity:	19.35%	Mismatches:	142
Query Match:	7.65%	Indels:	117
DB:	8	Gaps:	16

US-09-813-329-6 (1-409) x SCD8035 (1-69009)

Qy	51	ValAlaIleLeuAlaLeuThrIleTrpGlnThrArgValSerHisLeuAspLysGlu	70
Db	6086	GTGCTGTGTGAAGCTCTAAACACATTAAGTTCGACAGG-----	6124
Qy	71	LeuLysSerLeuLysArgValValAspAsnLeuGlnGlnArgLeuGlyIle-----	87
Db	6125	-----AACAAATTTTCTGACGACCTTCAAAATTTGAAGATCAT	6160
Qy	88	---AsnTyrLeuAspGluPheAspGluPheGlnLysGlu-----TyrGlu	101
Db	6161	GATGAAGTTGAAGACGAGGTTGACGAGCTCATAGGAAGAGATCCCGGATATGTTCAA	6220
Qy	102	AsnAlaLeuIleAspTyrProLysLysValAspGlyLeuThrAsp-----Glu	117
Db	6221	AACTCAGCGCTG-----GGATGTATGACGTAGAGGACATTTGAA	6259
Qy	118	GluAspAspAspAsp-----GlyAsp-----	124
Db	6260	GAAGACGACGATGACACACATCATTAATTGGACGCGATGCTATGGCATTTGTTGAT	6319
Qy	125	-----GlyLeuAspSerIleAlaAspAspGluAspAspAspValSerTyrSer	141
Db	6320	AGTGATAACGGATTGGAAGTAGTATTTAGTGACGAAGATGATGATGGGAGAGAAGAT	6379
Qy	142	ValAspAspValGlyAlaAspTyrGluAspTyrThrAspMetLeuAsnLysLeuAsnAsn	161
Db	6380	GCAGATGATGCTCGTTCAGATTCGGAGAA---AATGAATTATCTTCTGAATGCAATCA	6436
Qy	162	AlaHisThrGlyThrProThrSerGluThrThrAlaGluGlyGluGlyGluThrAsp	181
Db	6437	AGTACAGCGGACGGCAGCGATGTCGACTACGAGTTGATGATGCGAGTGGACTAATTATC	6496
Qy	182	SerAlaSerSerAlaSerAsnAspAsnValPheAspAspPhe-----ThrSer	198
Db	6497	AATATTGACCCGCCCTCAGGAGATGATGAAGAAATGGCAGACTATGATGCTTAACATTAGC	6556
Qy	199	TyrAsnAlaHisLysLysGlnGluArgLysSerArgSerIleAlaAspValArgAsn	218
Db	6557	CATTCTTCTACAGTGAANAATAGATGACCGCTCAATGGACGTCATAGAAGTTTACGAC	6616
Qy	219	GluGluGlnAsnIleGlnGlnAsnHisThrGluLeuGlnGlu-----	232
Db	6617	GACGAGCTAAGT---TCTGGATATGATGTTGATTTAACCGATTATGACGTAGATGAATCT	6673
Qy	233	-----LysSerSerAsnGluAlaThrSerLysGluSer	243
Db	6674	GATTGGGACTCCGGGTTATCCAGCCTATCTATTCTGATGAAGATAGTGAAGTAGCGAA	6733

Search completed: September 15, 2004, 10:23:34

Job time : 7208 secs

Qy	244	ProAlaProLeuHisHisArgArgMetHisSerArgHisArgHisLeuLeuValArg	263
Db	6734	GATGAGCTATTAACTCCACCAGATCGGAGATTCAGAAGGAGATCGCTGATAGTGAA	6793
Qy	264	LysAlaArg---SerGluAspSerArgProAlaAlaHisPheHisLeuSerSerArgArg	282
Db	6794	GGCGTGGAGCTAACTGACGACTCGCAA-----	6820
Qy	283	ArgHisGlnGlySerMetGlyTyrHisGlyAspMetTyrIleGlyAsnAspAsnGluArg	302
Db	6821	-----GGTGAGTGAAGAAGAT	6838
Qy	303	AsnSerTyrGlnGlyHisPheGlnThrArgAspGlyValLeuThrValThrAsnThrGly	322
Db	6839	GAT-----AGAGGTGTTTCCGTGGTATTGAACATATATTTTCAAAATGAGAATGAGCCT	6892
Qy	323	LeuTyrTyrValTyrAlaGlnIleCysTyrAsnAsnSerHisAspGlnAsnGlyPheIle	342
Db	6893	TTGTTTAGAGTTTCATGACGAAATCGCTCATAGAAACCATCATCGCTCAATCAATCGTACC	6952
Qy	343	ValPheGlnGly-----AspThrProPheLeuGlnCysLeuAsnThrValProThrAsn	360
Db	6953	CATTTTCACTCTGCTATGAGCGGCCCTCTTTAAGCTTATTAAATCTGTGAAGAAGAAAC	7012
Qy	361	MetProHisLysValHisThrCysHisThrSerGlyLeuIleHisLeuGlu-----	377
Db	7013	CAAAGCAACCTAATAATCCATAGGCTACAGGATTGGACAACAGTGGAGAACGATATT	7072
Qy	378	-----ArgAsnGluArgIleHisLeu	384
Db	7073	TCAGATCAGGTGACTGTTGCTGTTGAGGCTCTCGCCCAAGATCACATCATCTTCACTTT	7132
Qy	385	LysAspIle	387
Db	7133	TCGGAAGTT	7141



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: September 15, 2004, 02:17:51 ; Search time 634 seconds  
(without alignments)  
2740.557 Million cell updates/sec

Title: US-09-813-329-6  
Perfect score: 2162  
Sequence: 1 MTAETLKBFFITPTGANDGF.....DRNAVLRGNRSYFGIFKV 409

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n.model -DEV=xlp  
-Q/cgn2\_1/USPTO spool\_p/US09813329/runat\_13092004.102752.21656/app\_query.fasta\_1.583  
-DB=N Geneseq 29Jan04 -QWt=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09813329 @CGN 1.1.885 @runat\_13092004.102752.21656 -NCFU=6 -ICPU=3  
-NO WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N Geneseq 29Jan04:.\*  
1: geneseqn1980s:.\*  
2: geneseqn1990s:.\*  
3: geneseqn2000s:.\*  
4: geneseqn2001as:.\*  
5: geneseqn2001bs:.\*  
6: geneseqn2002as:.\*  
7: geneseqn2003as:.\*  
8: geneseqn2003bs:.\*  
9: geneseqn2003cs:.\*  
10: geneseqn2004s:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2162	100.0	2148	6 ABK11680	Abk11680 DNA encod
2	2082.5	96.3	1221	6 ABK11679	Abk11679 DNA encod
3	2045	94.6	2166	6 ABK11678	Abk11678 DNA encod
4	1656	76.6	978	4 ABL21473	AbL21473 Drosophil
5	1483.5	68.6	3324	4 ABL21472	AbL21472 Drosophil
6	151	7.0	2616	6 ABL54139	AbS54139 Mouse cDN
7	145.5	6.7	11528	4 ABL15995	AbL15995 Drosophil
8	141	6.5	1287	6 ABQ76511	AbQ76511 C. albica

c	9	140	6.5	567	3	AAA29550	Aaa29550 HIV codon
	10	138.5	6.4	3579	3	AAA70099	Aaa70099 Plasmodiu
	11	138.5	6.4	6240	4	ABL06443	AbL06443 Drosophil
	12	138	6.4	1176	6	ABL51020	AbL51020 Human EDA
	13	138	6.4	1176	8	ACD07906	AcD07906 DNA encod
	14	138	6.4	1574	6	ABL51009	AbL51009 Human EDA
	15	138	6.4	1574	8	ACD07895	AcD07895 Polynucle
	16	138	6.4	5307	7	ACC57577	Acc57577 Polynucle
	17	138	6.4	5307	7	ACC57910	Acc57910 Human EDA
	18	138	6.4	5307	9	ADC35223	AdC35223 Human EDA
	19	137	6.3	1357	3	AAZ51257	Aaz51257 Human RNA
	20	137	6.3	2635	4	ABL20713	AbL20713 Drosophil
c	21	137	6.3	4183	4	AAK74968	Aak74968 Human imm
c	22	137	6.3	4183	4	AAK75364	Aak75364 Human imm
	23	137	6.3	5276	4	ABL20712	AbL20712 Drosophil
	24	137	6.3	6796	4	ABL17763	AbL17763 Drosophil
	25	136	6.3	40059	4	ABL15994	AbL15994 Drosophil
	26	135	6.2	1626	4	AAS35027	Aas35027 cDNA enco
	27	135	6.2	1626	4	ABK43159	AbK43159 cDNA enco
	28	135	6.2	1626	9	ADC46185	AdC46185 Human neo
	29	135	6.2	2075	4	ABL21925	AbL21925 Drosophil
c	30	135	6.2	4248	4	ABL21924	AbL21924 Drosophil
	31	134	6.2	1401	7	ACC60802	Acc60802 Gene sequ
	32	133	6.2	2614	4	AAH16318	Aah16318 Human cDN
	33	133	6.2	3489	3	AAA30290	Aaa30290 Kaposi's
	34	133	6.2	3489	4	AAF82901	Aaf82901 Nucleotid
	35	133	6.2	3489	6	ABA93487	Aba93487 Kaposi's
	36	133	6.2	7503	3	AAW70206	Aaw70206 Plasmodiu
c	37	133	6.2	32207	2	AAV73805	Aav73805 KSHV LTR
c	38	133	6.2	137507	2	AAV19941	Aav19941 KSHV long
	39	130.5	6.0	3615	4	ABL03529	AbL03529 Drosophil
	40	129	6.0	4599	2	AAT27052	Aat27052 RRP3 telo
	41	128.5	5.9	1516	4	ABL17433	AbL17433 Drosophil
	42	128.5	5.9	3100	7	RAA51702	Raa51702 Human nuc
	43	128.5	5.9	3578	4	ABL17432	AbL17432 Drosophil
	44	128.5	5.9	4522	2	AAV74350	Aav74350 Staphyloc
	45	128	5.9	2619	6	ABZ32134	AbZ32134 Candida a

ALIGNMENTS

RESULT 1	ABK11680	ABK11680 standard; DNA; 2148 BP.
ID	ABK11680	
XX	XX	ABK11680;
AC	ABK11680;	
XX	XX	05-JUN-2002 (first entry)
DT	05-JUN-2002	
XX	XX	DNA encoding tumour necrosis factor variant 2 (TNFv2).
DE	DE	Tumour necrosis factor; TNF; pesticide; crop-damaging insect;
KW	KW	farm animal insect; epithelial morphogenesis; cell-matrix adhesion;
KW	KW	x-linked anhidrotic (hypohidrotic) ectodermal dysplasia;
KW	KW	x-linked anhidrotic (hypohidrotic) ectodermal dysplasia-like disorder;
KW	KW	sparse hair; sweat gland aberration; endotoxin shock; inflammation;
KW	KW	haemorrhagic necrosis of tumour; cytotoxicity; TNFv2;
KW	KW	obesity-linked insulin resistance; gene; ds.
XX	OS	Drosophila melanogaster.
XX	XX	
XX	XX	Key Location/Qualifiers
XX	XX	CDS 634..1863
XX	XX	/*tag= b
XX	XX	/product= "TNFv1"
XX	XX	/note= "Tumour necrosis factor variant 1"
XX	XX	misc_difference 634..1860
XX	XX	/*tag= d
XX	XX	/note= "Specifically claimed in claim 22"
XX	XX	sig_peptide 634..789
XX	XX	/*tag= a
XX	XX	mat_peptide 790..1860
XX	XX	/*tag= c

FT XX /label= mature\_TNFv1  
 PN US2002012968-A1.  
 XX 31-JAN-2002.  
 PD 20-MAR-2001; 2001US-00813329.  
 PF XX  
 XX 21-MAR-2000; 2000US-0190816P.  
 PR (CARR/) CARROLL P M.  
 XX (CHEN/) CHEN J.  
 PA (RAMA/) RAMANATHAN C S.  
 PA (XIAO/) XIAO H.  
 PA (GUAN/) GUAN B.  
 PA (BOWE/) BOWEN M A.  
 XX  
 PI Carroll PM, Chen J, Ramanathan CS, Xiao H, Guan B, Bowen MA;  
 XX  
 DR WPI: 2002-195121/25.  
 DR P-PSDB; AAU77718.  
 DR  
 XX New Drosophila tumor necrosis factor molecule, useful in controlling  
 PT agriculturally important pests, e.g. comprises modifying the growth,  
 PT feeding or reproduction of crop-damaging insects or insects of farm  
 PT animals.  
 XX  
 PS Claim 2; Fig 3A-C; 11pp; English.  
 XX  
 CC The invention describes an isolated tumour necrosis factor polypeptide  
 CC (TNF). The polypeptide and polynucleotide are useful in controlling  
 CC agriculturally important pests, particularly by modifying the growth,  
 CC feeding and/or reproduction of crop-damaging insects or insects of farm  
 CC animals. The polypeptide and polynucleotide are useful for modulating  
 CC epithelial morphogenesis, cell-matrix adhesion in flies and mammals. Thus  
 CC the polypeptide and polynucleotide may be useful for treating,  
 CC ameliorating or preventing X-linked anhidrotic (hypohidrotic) ectodermal  
 CC dysplasia and X-linked anhidrotic (hypohidrotic) ectodermal dysplasia-  
 CC like disorders, e.g. sparse hair, abnormal or missing teeth or sweat  
 CC gland aberrations in animals (e.g. insects and potentially humans),  
 CC endotoxic shock, inflammation, hemorrhagic necrosis of tumours,  
 CC cytotoxicity and obesity-linked insulin resistance, all of which involve  
 CC TNF molecules. This sequence encodes the drosophila melanogaster tumour  
 CC necrosis factor variant 2 (TNFv2) protein, described in the invention  
 XX  
 SQ Sequence 2148 BP; 645 A; 510 C; 506 G; 487 T; 0 U; 0 Other;  
 Alignment Scores:  
 Pred. No.: 3.52e-179 Length: 2148  
 Score: 2162.00 Matches: 409  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 6 Gaps: 0  
 US-09-813-329-6 (1-409) x ABK11680 (1-2148)  
 Qy 1 MetThrAlaGluThrLeuLysProPheIleThrProThrSerAlaAsnAspGlyPhe 20  
 Db 634 ATGACTGCGGAGACCTTCAGCCGTTTATACGCCAACGAGTGCCACGATGATGTTT 693  
 Qy 21 ProAlaLysAlaThrSerThrAlaThrAlaGlnArgArgThrArgGlnLeuIleProLeu 40  
 Db 694 CCGGCCAAAGCGACAGCGGCGACCGCCGCGAGCGACCGCCGAGCTGATCCCGCTG 753  
 Qy 41 ValLeuGlyPheIleGlyLeuGlyValValAlaIleLeuAlaIleThrIleTrpGln 60  
 Db 754 GTTTTGGGGTTCATCGGTCTGGGGCTGGTCTGGCATTCGCACTAAGCACTCTGGCAG 813  
 Qy 61 ThrThrArgValSerHisLeuAspLysGluLeuLysSerLeuLysArgValValAspAsn 80  
 Db 814 ACAACGGTGTATCGCATCTGGACAGGAGCTGAAGAGCGCTGAAGCGAGTCTCGTAAAT 873

Qy 81 LeuGlnGlnArgLeuGlyIleAsnTyrLeuAspGluPheAspGluPheGlnLysGluTyr 100  
 Db 874 CTCAGCAGCGCTTGGGCATAAACTATCTGGACGAGTTCGACGAGTTCCAAAGAGGAGTAC 933  
 Qy 101 GluAsnAlaLeuIleAspTyrProLysLysValAspGlyLeuThrAspGluAspAsp 120  
 Db 934 GAGAAATGCCCTCATCGACTATCCAAAAAGGTGGATGGCTTCACGATGAGGAGGACGAC 993  
 Qy 121 AspAspGlyAspGlyLeuAspSerIleAlaAspAspGluAspAspValSerTyrSer 140  
 Db 994 GACGATGGCGATGGTCTGGATTCCATTGGGACGACGAGGACGACGATTTAGCTATAGC 1053  
 Qy 141 SerValAspAspValGlyAlaAspTyrGluAspTyrThrAspMetLeuAsnLysLeuAsn 160  
 Db 1054 TCTGTGATGATGTTGGCGCAGACTACGAGGACTACACCGATATGTTAAATAAATCAAC 1113  
 Qy 161 AsnAlaHisThrGlyThrProThrSerGluThrThrAlaGluGlyGluGlyThr 180  
 Db 1114 AATGCACATACCGGCACCGCCACATCTGAGACCATCTGAGGGCGGCGAGGCGAGCG 1173  
 Qy 181 AspSerAlaSerSerAlaSerAsnAspAsnValPheAspAspPheThrSerTyrAsn 200  
 Db 1174 GACAGTGCATCTCTCAGCCTCAATGATGACATGTTGTCGATGACTTTACCGAGCTACAAT 1233  
 Qy 201 AlaHisLysLysLysGlnGluArgLysSerArgSerIleAlaAspValArgAsnGlu 220  
 Db 1234 GCCCACAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1293  
 Qy 221 GlnAsnIleGlnGlyAsnHisThrGluLeuGlnGlyLysSerSerAsnGluAlaThrSer 240  
 Db 1294 CAGATATTTCAGGAAATCACAAGAGCTTCAGGAAAGTCAATCCAAATGAGGCACTTCC 1353  
 Qy 241 LysGluSerProAlaProLeuHisArgArgArgMetHisSerArgHisArgHisLeu 260  
 Db 1354 AAAGAGAGCGCTGCACCACTTCCACCGTGCAGATGATGCTCCGCGCATCGCCACCTC 1413  
 Qy 261 LeuValArgLysAlaArgSerGluAspSerArgProAlaAlaHisPheHisLeuSerSer 280  
 Db 1414 CTAGTCCGCAAGCCAGATCCGAGGACTCGAGGCGCAGGAGGAGGAGGAGGAGGAGGAG 1473  
 Qy 281 ArgArgArgHisGlnGlySerMetGlyTyrHisGlyAspMetTyrIleGlyAsnAspAsn 300  
 Db 1474 AGCGCGCGCTCCCAAGGAGATGCGCTACCATGGAGATATGTACATAGGAAATGATTAAC 1533  
 Qy 301 GluArgAsnSerTyrGlnGlyHisPheGlnThrArgAspGlyValLeuThrValThrAsn 320  
 Db 1534 GAGAGAAATCTTTATCAGGAGACCTTTCAAACGCGCGATGGCGTCTTGACGCTGACCAAT 1593  
 Qy 321 ThrGlyLeuTyrTyrValTyrAlaGlnIleCysTyrAsnAsnSerHisAspGlnAsnGly 340  
 Db 1594 ACAGGCGCTATATTACGTATACGCCAGATATGCTACAACTCGACGAGGAGGAGGAGGAG 1653  
 Qy 341 PheIleValPheGlnGlyAspThrProPheLeuGlnCysLeuAsnThrValProThrAsn 360  
 Db 1654 TTTATCGTCTTTCAAGGAGACACTCCATTCTCGAGTGTCTTGAACCGGTGCCCAACCAAC 1713  
 Qy 361 MetProHisLysValHisThrCysHisThrSerGlyLeuIleHisLeuGluArgAsnGlu 380  
 Db 1714 ATGCCACATAAGGAGTGACACCTGCCACACGATGCTGATCTCCACCTGGAAACGAAACGAG 1773  
 Qy 381 ArgIleHisLeuLysAspIleHisAsnAspArgAsnAlaValLeuArgGluGlyAsnAsn 400  
 Db 1774 AGGATCCATCTGAAGGACATTCACACGATCGCAATGCAGTTCGAGTTCGCGGAGGAGGAG 1833  
 Qy 401 ArgSerTyrPheGlyIlePheLysVal 409  
 Db 1834 CGAAGACTACTTTGGCATCTTCAAGGTG 1860  
 RESULT 2  
 ID ABK11679 standard; DNA; 1221 BP.  
 XX  
 AC ABK11679;

XX 05-JUN-2002 (first entry)  
 XX DNA encoding tumour necrosis factor variant 1 (TNFv1).  
 DE  
 XX Tumour necrosis factor; TNF; pesticide; crop-damaging insect;  
 KW farm animal insect; epithelial morphogenesis; cell-matrix adhesion;  
 KW X-linked anhidrotic (hypohidrotic) ectodermal dysplasia;  
 KW X-linked anhidrotic (hypohidrotic) ectodermal dysplasia-like disorder;  
 KW sparse hair; sweat gland aberration; endotoxigenic shock; inflammation;  
 KW haemorrhagic necrosis of tumour; cytotoxicity; TNFv1;  
 KW obesity-linked insulin resistance; gene; ds.  
 XX Drosophila melanogaster.  
 OS  
 XX Key Location/Qualifiers  
 FH CDS 1..1221  
 FT /\*tag= b  
 FT /product= "TNFv1"  
 FT /note= "Tumour necrosis factor variant 1"  
 FT misc\_difference 1..1218  
 FT /\*tag= d  
 FT /note= "Specifically claimed in claim 18"  
 FT sig\_peptide 1..156  
 FT /\*tag= a  
 FT mat\_peptide 157..1218  
 FT /\*tag= c  
 FT /label= mature\_TNFv1  
 XX  
 PN US2002012968-A1.  
 XX  
 XX 31-JAN-2002.  
 PD  
 XX 20-MAR-2001; 2001US-00813329.  
 PF  
 XX 21-MAR-2000; 2000US-0190816P.  
 PR  
 XX (CARR/) CARROLL P M.  
 PA (CHEN/) CHEN J.  
 PA (RAMA/) RAMANATHAN C S.  
 PA (XIAO/) XIAO H.  
 PA (GUAN/) GUAN B.  
 PA (BOWE/) BOWEN M A.  
 XX  
 XX Carroll PM, Chen J, Ramanathan CS, Xiao H, Guan B, Bowen MA;  
 PI WPI; 2002-195121/25.  
 XX  
 XX New Drosophila tumor necrosis factor molecule, useful in controlling  
 PT agriculturally important pests, e.g. comprises modifying the growth,  
 PT feeding or reproduction of crop-damaging insects or insects of farm  
 PT animals.  
 XX  
 PS Claim 2; Fig 2A-B; 119pp; English.  
 XX  
 XX The invention describes an isolated tumour necrosis factor polypeptide  
 CC (TNF). The polypeptide and polynucleotide are useful in controlling  
 CC agriculturally important pests, particularly by modifying the growth,  
 CC feeding and/or reproduction of crop-damaging insects or insects of farm  
 CC animals. The polypeptide and polynucleotide are useful for modulating  
 CC epithelial morphogenesis, cell-matrix adhesion in flies and mammals. Thus  
 CC the polypeptide and polynucleotide may be useful for treating,  
 CC ameliorating or preventing X-linked anhidrotic (hypohidrotic) ectodermal  
 CC dysplasia and X-linked anhidrotic (hypohidrotic) ectodermal dysplasia-  
 CC like disorders, e.g. sparse hair, abnormal or missing teeth or sweat  
 CC gland aberrations in animals (e.g. insects and potentially humans),  
 CC endotoxigenic shock, inflammation, haemorrhagic necrosis of tumours,  
 CC cytotoxicity and obesity-linked insulin resistance, all of which involve  
 CC TNF molecules. This sequence encodes the drosophila melanogaster tumour  
 CC necrosis factor variant 1 (TNFv1) protein, described in the invention  
 XX  
 XX Sequence 1221 BP; 349 A; 323 C; 312 G; 237 T; 0 U; 0 Other;  
 SQ

Alignment Scores:  
 Pred. No.: 1.54e-172 Length: 1221  
 Score: 2082.50 Matches: 400  
 Percent Similarity: 96.39% Conservatives: 0  
 Best Local Similarity: 96.39% Mismatches: 0  
 Query Match: 96.32% Indels: 15  
 DB: 6 Gaps: 2  
 US-09-813-329-6 (1-409) x ABK11679 (1-1221)  
 Qy 1 MetThrAlaGluThrLeuLysProPheIleThrProThrSerAlaAanAspAspGlyPhe 20  
 Db 1 ATGACTCCGAGACCCCTCAAGCCGTTTATACGCCAACGAGTGCACACGATGATGGTTTT 60  
 Qy 21 ProAlaLysAlaThrSerThrAlaThrAlaGlnArgThrArgGlnLeuLeuProLeu 40  
 Db 61 CCGGCCAAGGACGACGACGCGCCGAGGACGACCCGCGAGCTGATCCCCCTG 120  
 Qy 41 ValLeuGlyPheLeGlyLeuValValAlaLeuAlaLeuLeuLeuLeuLeuLeuLeu 60  
 Db 121 GTTTTGGGGTTTCATCGGCTCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180  
 Qy 61 ThrThrArgValSerHisLeuAspLysLeuLysLeuLysSerLeuLysArgValValAspAsn 80  
 Db 181 ACAACGCGTGTATCGCATCTGGACAAGGAGCTGAAGAGCCTGAAGCGAGTCTGCGATAAT 240  
 Qy 81 LeuGlnGlnArgLeuGlyIleAsnTyrLeuAspGluPheAspGluPheGlnLysGluTyr 100  
 Db 241 CTCGAGCAGCGTTTGGGCNTAACTATCTGGACGAGTTCGACGAGTTCCAAAGGAGGTAC 300  
 Qy 101 GluAsnAlaLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 120  
 Db 301 GAGAAATGCCCTCATCGACTATCCAAAAAAGGTGGTGGCTCTACGGATGAGGAGGACGAC 360  
 Qy 121 AspAspGlyAspGlyLeuAspSerIleAlaAspAspGluAspAspAspValSerTyrSer 140  
 Db 361 GACGATGCGATGGTCTGGATTCATTTGCCGACGACGAGGACGACGACGCTAGGTATAGC 420  
 Qy 141 SerValAspAspValGlyAlaAspTyrGluAspTyrThrAspMetLeuAsnLysLeuAsn 160  
 Db 421 TCTGTGGATGATGTTGGCGCAGACTACGAGGACTACACCCGATATGTTAAATAAATCAAC 480  
 Qy 161 AsnAlaHisThrGlyThrThrProThrSerGluThrThrAlaGluGlyGluGlyGluThr 180  
 Db 481 AATGCACATACCGCACACCGCCACATCTGAGACCACTCTGAGGCGGAGGCGGAGGAG 540  
 Qy 181 AspSerAlaSerSerAlaSerAsnAspAspAsnValPheAspAspPheThrSerTyrAsn 200  
 Db 541 GACAGTGCATCTCTCAGCCTCAAAATGATGACAATGTGTTCGATGACTTTACGAGTACAAT 600  
 Qy 201 AlaHisLysLysLysGlnGluArgLysSerArgSerIleAlaAspValArgAsnGluGlu 220  
 Db 601 GCCCAAAAAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 660  
 Qy 221 GlnAsnIleGlnGlyAsnHisThrGluLeuGlnGlnLysSerSerAsnGluAlaThrSer 240  
 Db 661 CAGAAATTTCAAGAAATACACAGAGCTTCAGGAAAGTATCCCAATGAGGCAACTTCC 720  
 Qy 241 LysGluSerProAlaProLeuHisHisArgArgArgMetHisSerArgHisArgHisLeu 260  
 Db 721 AAAAG-----AGAAATGCAATTTCCGCCCATCGCACCTC 753  
 Qy 261 LeuValArgLys-----AlaArgSerGluAspSerArgProAlaAla 274  
 Db 754 CTAGTCCGCAAGGTGAATCTCTTCTTTCAGCCGATCCGAGGACTCCAGGCGCAGCAGCC 813  
 Qy 275 HisPheHisLeuSerSerArgArgArgHisGlnGlySerMetGlyTyrHisGlyAspMet 294  
 Db 814 CATTTCCACTTGAGCAGCGCGCGTCCACCAAGGAAGTATGGCTACCATGGAGATATG 873  
 Qy 295 TyrIleGlyAsnAspAsnGluArgAsnSerTyrGlnGlyHisPheGlnThrArgAspGly 314  
 Db 874 TACATAGGAATGATACGAGAGAAACTCTTATCAGGACACTTTTCAACGCGCGATGGC 933



Qy	181	AspSerAlaSerSerAlaSerAsnAspAsnValPheAspAspPheThrSerTyrAsn	200
Db	1192	GACAGTGCATCCTCAGCCTCAAAATGATGACAATGTGTTTCGATGACTTTTACACAGCTCGAAT	1251
Qy	201	AlaHisLeuLysLysGlnGluArgLysSerArgSerIleAlaAspValArgAsnGluGlu	220
Db	1252	GCCCTCAAAAAGAGCAGGAGAGAAAATCTCGCTCGATTCGGATGTACGCAATGAGGAG	1311
Qy	221	GlnAsnIleGlnGlyAsnHisThrGluLeuGlnGluLysSerSerAsnGluAlaThrSer	240
Db	1312	CAGAATATTCAAGGAAATCACACAGAGCTTCAGAGAAAAGTCATCCAATGAGCGAGCTTCC	1371
Qy	241	LysGluSerProAlaProLeuHisHisArgArgMetHisSerArgHisArgHisLeu	260
Db	1372	AAAGAGAGCCCTGCAGCACTTCACCTCGTTCGAGNATGCATTTCCGCCATCGCACCTC	1431
Qy	261	LeuValArgLysAlaArgSerGluAspSerArgProAlaAlaHisPheHisLeuSerSer	280
Db	1432	GTAGTCGCGAAAGCCAGATCCGAGGACTTCGAGGCGCAGCAGGCCCATTTTCCACTTGAGCAGC	1491
Qy	281	ArgArgArgHisGlnGlySerMetGlyTyrHisGlyAspMetTyrIleGlyAsnAspAsn	300
Db	1492	AGCGCGGTACCAAAGAAAGTATGGGCTACCATCGAGATATGTACATAGAAAATGATAGG	1551
Qy	301	GluArgAsnSerTyrGlnGlyHisPheGlnThrArgAspGlyValLeuThrValThrAsn	320
Db	1552	GAGAGATGCTCTTATCAGGACACTTTCAAACGCGCATGTCGCTATTGACGGTGACCAAT	1611
Qy	321	ThrGlyLeuTyrTyrValTyrAlaGlnIleCysTyrAsnAsnSerHisAspGluAsnGly	340
Db	1612	CGAGGCTTATTATCAGTATACGCCCAAGATATGGGGCTCAAACTCGCAGCAGACCAAGCGGA	1671
Qy	341	PheIleValPheGlnGlyAspThrProPheLeuGlnCysLeuAsnThrValProThrAsn	360
Db	1672	TTTATCGTCTTCAAGGAGACACTCCATTCTCTGAGTGTCTTGAACACGGTGCCCAACAC	1731
Qy	361	MetProHisLysValHisThrCysHisThrSerGlyLeuIleHisLeuGluArgAsnGlu	380
Db	1732	ATGCCACATAGGTGTCACCTGCGCACAGTGGTCTGATTCACCTCGGAACGAAAGGAG	1791
Qy	381	ArgIleHisLeuLysAspIleHisAsnAspArgAsnAlaValLeuArgGluGlyAsnAsn	400
Db	1792	AGGATCCATCTGAAGGACATTCACACGATCCCAATCCGATTCGCGGAGGGAAAACAC	1851
Qy	401	ArgSerTyrPheGlyIlePheLysVal	409
Db	1852	CGAAGCTACTTTGGCATCTCTCAAGGTG	1878

RESULT 4  
ABL21473  
ID ABL21473 standard; DNA: 978 BP.

DT 26-MAR-2002 (first entry)

DE *Drosophila melanogaster* genomic polynucleotide SEQ ID NO 15892.

**KW** Drosophila; developmental biology; cell signalling; insecticide; pharmaceutical; gene; ds.

OS *Drosophila melanogaster*.

PN WO200171042-A2.

PD 27-SEP-2001.

23-MAR-2001: 2001WO-US009231.

PR 23-MAR-2000; 2000US-0191637P.

XX  
44-38861-1000

Venter JC, Adams M, Li PWD, Myers EW;  
WPI; 2001-656860/75.

New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signaling and cell-cell interactions.

Claim 1; SEQ ID NO 15892; 21pp + Sequence Listing; English.

The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (AB116176-ABL30511), expressed DNA sequences (AB101840-ABL16175) and the encoded proteins (AB857737-AB872072). The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIFO at [ftp.wifo.int/pub/published\\_pct\\_sequences](http://ftp.wifo.int/pub/published_pct_sequences)

Sequence 978 BP; 289 A; 261 C; 238 G; 130 T; 0 U; 0 Other;

PA





Db 1402 CAGGCAACCTAATGAAGTCACCTTTCTGAATTTCTCA -GGTACCATCGAGATATGTACAT 1344

Qy 296 eGlyAsnAspAsnGluArgAsnSerTyrGlnGlyHisPheGlnThrArgAspGlyValle 316

Db 1343 AGGAAATGATAACGAGAGAACTCTTATCAGGGACACTTTCNAAGCGCGATGGCGTCTT 1284

Qy 316 uThrValThrAsnThrGlyLeuTyrTyrValTyrAlaGlnIleCysTyrAsnAsnSerHi 336

Db 1283 GACGGTGACCAATACAGGCCTATATTAGGTATATACGCCAGATATGCTACAACTCGCA 1224

Qy 336 sAspGlnAsnGlyPheIleValPheGlnGlyAspThrProPheLeuGlnCysLeuAsnTh 356

Db 1223 GCACCAAGACGATTTATCGCTTTCAAGGAGACACTCCATCTCCGAGTGTCTGTAACAC 1164

Qy 356 rValProThrAsnMetProHisLysValHisThrCysHisThrSerGlyLeuIleHisLe 376

Db 1163 GGTGCCCAACCAATGCCACATAGTGCACACCTGCCACAGAGTGTCTGTATCCACCT 1104

Qy 376 uGluArgAsnGluArgIleHisLeuLysAspIleHisAsnAspArgAsnAlaValLeuAr 396

Db 1103 GGAACGAACGAGAGGATCCATCTGAAGGACATTCCACACGATCGCAATGCAGTCTCGC 1044

Qy 396 gGluGlyAsnAsnArgSerTyrPheGlyIlePheIysVal 409

Db 1043 GGAGGAAACCAACCGAAGTACTTTGGCATCTTCAAGGTG 1004

RESULT 6

ID ABS54139 standard; cDNA; 2616 BP.

XX AC ABS54139;

XX DT 25-NOV-2002 (first entry)

XX DE Mouse cDNA encoding Zis-SR protein.

XX KW Mouse; ss; gene; Zis-SR; neuroendocrine phenotype; diabetes;

XX KW Parkinson's disease; Alzheimer's disease; neurodegenerative disease;

XX KW zinc finger splicing with extended Ser-Arg domain; secretory pathway;

XX KW zinc finger protein.

OS Mus sp.

XX Key 23.1015 Location/Qualifiers

XX CDS /\*tag= a

XX FT /product= "Zis-SR"

XX PN WO200261082-A2.

XX PD 08-AUG-2002.

XX PF 29-JAN-2002; 2002WO-CA000101.

XX PR 29-JAN-2001; 2001US-0264296P.

XX PA (UYSH ) UNIV SHERBROOKE.

XX PI Day R;

XX DR WPI; 2002-682683/73.

XX DR P-ESDB; ABG32871.

XX PT New Zis-SR nucleic acid molecules and polypeptides, useful for restoring

XX PT or increasing the secretory properties of a cell, or for treating

XX PT diseases or conditions associated with a loss of function, e.g. diabetes

XX PT or Parkinson's disease.

XX PS Claim 1; Fig 4; 70pp; English.

XX CC The invention relates to an isolated nucleic acid molecule, Zis-SR,

XX CC encoding a protein involved in the secretory pathway in a cell ( or its

XX CC homologue or variant) or nucleic acid molecules that hybridise under high

stringency condition to the Zis-SR nucleic acid. Also included are an isolated polypeptide involved in the formation of secretory granules in cells comprising the amino acid sequence spanning amino acids 243-310 of the Zis-SR protein, restoring the neuroendocrine differentiation of a cell using the nucleic acid molecule or polypeptide cited above, identifying a gene and/or protein involved in inducing regulated secretion comprising a comparison at the molecular level of a secretion-defective cell line under conditions that restore differentiation of the secretion-defective cell, such that secretion is restored, and the secretion-defective cell line in the absence of the conditions cited. Also included are modulating the secretory properties of a cell comprising modulating the activity and/or level of Zis-SR and an assay to identify a modulator of regulated secretion in a cell comprising an assessment of a biological activity of Zis-SR, its part or derivative in the presence of a candidate agent, where a modulator of regulated secretion is selected when the biological activity of Zis-SR, its part or derivative is measurably different in the presence of the candidate compound as compared in its absence. The nucleic acid molecules or polypeptides are useful for restoring or increasing the secretory properties of a cell, for regulating neuroendocrine phenotype, and for long term therapies to treat diseases or conditions associated with a loss of function, e.g. diabetes, neurodegenerative diseases such as Alzheimer's disease or Parkinson's disease. The assay is useful for screening compounds for treating diseases or conditions associated with a defect in the regulated secretory pathways in cells. The nucleic acid molecules can also be used to locate gene regions associated with genetic diseases. The present sequence encodes mouse Zis-SR (zinc finger splicing with extended Ser-Arg domain)

XX SQ Sequence 2616 BP; 834 A; 458 C; 542 G; 782 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 0.00257 Length: 2616

Score: 151.00 Matches: 93

Percent Similarity: 33.84% Conservative: 63

Best Local Similarity: 20.17% Mismatches: 156

Query Match: 6.98% Indels: 150

DB: 6 Gaps: 17

US-09-813-329-6 (1-409) x ABS54139 (1-2616)

Qy 20 PheProAlaLysAlaThrSerThrAlaThrAlaGlnArgThrArgGlnLeuLeuPro 39

Db 95 TTTTGCTAGGAGAACCCAGCTGTAAACAGATGTGTGTCGAGAAAAGACAACTGAGGCCAAGATG 154

Qy 40 LeuValLeuGlyPheIleGlyLeuGlyLeuValValAla-----IleLeu 54

Db 155 ATGAAAGCTGGGGAAACAGAAATAGGAAAGACACTGGCAGAGAGAGCCGGGCTTATTT 214

Qy 55 AlaLeuThrIleTrpGlnThrThrArgValSerHisLeuAspLysGluLeuLysSerLeu 74

Db 215 AGTGCCCAATGATTGGCAATGCCAAACTTGCAGTAATGTGAATGGGCTAGAGATCAGAG 274

Qy 75 LysArgValValAsp-----AsnLeuGlnGlnArgLeuGly----- 86

Db 275 TGTAAACATGTGTAATACTCCAAAGATGTCTAAATTTAGAGAAAGAACAGGATATGGAGGT 334

Qy 87 -----IleAsnTyrLeuAsp-----GluPhe 93

Db 335 GGTTTAAATGAAAGAGAGAATGTTGAATACATAGAAGAGAGAAGAAATCTGTATGGGAATAT 394

Qy 94 AspGluPheGlnLysGlyTyrGluAsnAlaLeuIleAspTyrProLysLysValAspGly 113

Db 395 GATGAGTTGGAGCTGTAAGAAAAGAAAA-----TACAGGGGGAAGCAGTGGC 442

Qy 114 -----LeuThrAspGluGluAspAspAspGlyAspGlyLeuAspSerIleAla 130

Db 443 CCTGCATCTATATTAAAGGAAGTTGAAGATAAAGAGTCAGAGGAGAGAA----- 493

Qy 131 AspAspGluAspAspAspValSerTyrSerSerValAspAspValGlyAlaAspTyrGlu 150

Db 494 GAGGATGAGGATGAAGATCTTCTTAAGTATTAAGCTAGAT----- 532

151	AspTyrThrAspMetLeuAnsnLysLeuAenAsnAlaHisThrGlyThrThrProThrSer	170
532	-----	532
171	GluThrThrAlaGluGlyGluGlyGluThrAspSerAlaSerSerAlaSerAsnAspAsp	190
533	-----GAGACGAGGATGAAGATGATGCTGATCTCTCAAAATATATAT	574
191	AsnValPheAspAspPheThrSerTyrAenAlaHisLysLysLysGlnGluArgLysSer	210
575	-----CTTGATGCCACGCAAGAGATAGTAAACAAAGAAAGCAATAGGCGGAGC	628
211	ArgSerIleAlaAspValArgAsnGluGluGlnAsnIleGlnGlyAsnHisThrGluLeu	230
629	CGCTCA-----	634
231	GlnGluLysSerSerAsnGluAlaThrSerLysGluSerProAlaProLeuHisArg	250
635	ANGTCACGATCGTCTCACTCAAGGCTTCATCACGCTCATCTCCCTCTCAAGTTCAAGG	694
251	ArgArgMetHisSerArgHisArgHisLeuLeuValArgLysAlaArgSer	267
695	TCCAGGTCAGGTCCTCCGTTCAAGAAGCTTCCAGTTCGCAGTCGAGGTCCTCACTCCCGT	754
268	-----GluAspSerArgProAlaAlaHisPheHisLeuSerSerArgArgHisGln	285
755	TCCAGAGAACATTCAGATCCCGTGGTTCGAAATCAAGATCCAGCTCCAGGTCCTCCACAG	814
286	GlySerMetGlyTyrHisGlyAspMetTyrIleGlyAsn-----Asp	299
815	GGCTCTCTTCCCAAGAAAAAGATCTTATTTCGAGTTCTTCGTCAATCTCTGAAAGAGAC	874
300	AsnGluArgAsnSerTyrGlnGlyHisPhe-----GlnThrArgAspGlyVal	315
875	AGGAAGAGAGTCGCTCTAGACC-TTCTTTCACCAGCTGTTTCGCAAAAAAGACGACGAG	933
316	LeuThrValThrAsnThrGlyLeuTyrTyrValTyr-----	327
934	ATCAGCGGTCAACCCGAAAGGCACCAAGGTCGCTCTCCGGATCAACACATTCCTGGTTCCCG	993
328	-----AlaGlnIleCysTyrAsnAsnSerHis-----	336
994	TTCAAGTTCAAAAAAGAAATAATGTATTAAAAATTTCATCTTAAAAAACATTGAGTACA	1053
336	-----	336
1054	GTGCATGAAGCATATTTTTAGGAAGTTGATGCTCATCTTTCGTGAGAGTACTACATCTG	1113
337	-----AspGlnAsnGlyPheIleValPheGlnGlyAspThrProPheLeuGlnCysLeu	354
1114	CTAGTAGAGGTGCATGCTTTATTGCTTTTCAAACAATAACAACGTGTTTATTGT---	1170
355	AsnThrValProThrAsnMetProHisLysValHisThrCysHisThrSerGlyLeuIle	374
1171	---GAAACTGAAGTAAATTGCATTTTAAAGCCATATATCTGTTCCTCCAGATAAATGCTCTG	1227
375	HisLeuGluArgAsnGluArgIleHisLeuLysAspIleHisAsnAspArgAsnAlaVal	394
1228	TTCAATTATTACACCACTTCCTTCATTTAAAA-----CATCTAGCAATAACAGAGTA	1281
395	Leu 395	
1282	CTT 1284	

Qy	126	LeuAspSerIleAlaAspGluAspAspValSerTyrSerSerValAspAspVal	145
Db	1574	-----GATGATGATGATGACGAC-----GACGACATGGAGCTGCTG	1609
Qy	146	GlyAlaAspTyrGluAspTyrThrAspMet-----LeuAsnLysLeuAsnAsnAlaHis	163
Db	1610	AGTGCAGTGATGATCATCTTCAACCCCTGTGCGCACTAAATGACACCAACAGCGCTG	1669
Qy	164	ThrGlyThrThrProThrSerGluThrThrAlaGluGlyGluGlyGluThrAspSerAla	183

Qy	126	LeuAspSerIle	ala	asp	asp	glu	asp	asp	asp	val	ser	tyr	ser	ser	val	asp	asp	145
Db	1574	---	-	GAT	GAT	GAT	GAT	GAT	GAT	GAT	GAT	GAT	GAT	GAT	GAT	GAT	GAT	1609
Qy	146	Gly	ala	asp	tyr	glu	asp	tyr	thr	asp	met	---	---	---	---	---	---	163
Db	1610	AGT	GCC	GAG	TGT	GAT	GAT	TTT	CA	CC	TGT	CGC	AA	CT	TA	AT	GA	1669
Qy	164	Gly	thr	thr	thr	pro	thr	ser	glu	thr	ala	glu	glu	glu	glu	thr	asp	183









CC endocrine sweat gland development in individuals from whom the normal  
CC sweating mechanism is compromised by disease or surgery, for stimulating  
CC the growth of mammary epithelial tissue, either for reconstructing or  
CC cosmetic purposes, and for promoting or maintaining differentiation of  
CC breast epithelium. (I) is also useful for functional analysis, antibody  
CC production and patient therapy. ABUS1009 to ABUS112 and ABUS0982 to  
CC ABUS0909 represent sequences used in the exemplification of the present  
CC invention



PT Increasing or decreasing hair follicle development, tooth development, or  
PT sweat gland development, in a tissue, involves modulating ectodermal  
PT dysplasia protein isoform II activity in the tissue.

PS Example 1; Page 59-60; 102pp; English.

XX The invention relates to a method of increasing or decreasing one or more  
CC of hair follicle development, tooth development, or sweat gland  
CC development, in a tissue, involves altering ectodermal dysplasia (EDAL)  
CC isoform II (EDAL-II) activity in the tissue. The method is useful for  
CC increasing or decreasing hair follicle development, tooth development or  
CC sweat gland development, in a tissue of a patient suffering from an  
CC ectodermal disease such as X-linked hypohidrotic ectodermal dysplasia  
CC (XLHED), autosomal recessive hypohidrotic ectodermal dysplasia (HED) or  
CC alopecia. The method is useful for decreasing hair follicle development  
CC in a subject suffering from hirsutism. The method is also useful for  
CC treating breast cancer and ectopic teeth, burns and trauma of skin due to  
CC surgery. The present sequence represents DNA encoding human EDAL-II

XX Sequence 1176 BP; 271 A; 354 C; 335 G; 216 T; 0 U; 0 Other;

# Alignment Scores:

Pred. No.: 0.0129 Length: 1176  
Score: 138.00 Matches: 96  
Percent Similarity: 35.92% Conservative: 57  
Best Local Similarity: 22.54% Mismatches: 147  
Query Match: 6.38% Indels: 127  
DB: 8 Gaps: 18

US-09-813-329-6 (1-409) x ACD07906 (1-1176)

QY 35 ArgGlnLeuIlePro-LeuValLeuGlyPheIleGlyLeuGlyLeuValValAlaIleLe 54  
DB 108 AGGGACAGCTGCTCTTCCTGGGTTCTTTGGCTCTCGCTGGCCCTCCACCTGCT 167  
QY 54 uAlaLeuThrIleTrpGlnThrThrArgValSerHisLeuAspLysGluLeuYssSerLe 74  
DB 168 GACGTTGTGCTGCTAC-----CTAGAGTTGGCTCGGA 200  
QY 74 uLysArgValValAspAsnLeuGlnArgLeuGlyValLeuAsnTyrLeuAspGluPheAs 94  
DB 201 GTTCGGCGGGAACGTGGAGCCGAGTCCGCTGGC----- 237  
QY 94 pGluPheGlnLysGluTyrGluAsnAlaLeuIleAspTyrProLysLysValAspGlyLe 114  
DB 237 ----- 237  
QY 114 uThrAspGluAspAspAspGlyAspGly-----LeuAs 127  
DB 238 -----GGCTCGGGCACCCCTGGCACCTCTGGCACCCCTAAG 272  
QY 127 pSerIleAlaAsp---AspGluAspAspAspValSerTyrSerSerValAspAspValGI 146  
DB 273 CAGCTCGGTGGCTCGACCTCGACGCCCATCACC----- 309  
QY 146 YAlaAspTyrGluAspTyrThrAspMetLeuAsnLysLeuAsnAlaHisThrGly-- 165  
DB 310 -----AGTCACTTGGGCA 323  
QY 166 -ThrThrProThrSerGluThrThrAlaGluGlyGluGlyGluThrAspSerAlaSerSe 185  
DB 324 GCCGTCACTAAGCAGCAGCATTTGGAACCGGGAGAGCCGCACTCCACTCTGACTCCCA 383  
QY 185 rAlaSerAsnAsp-----AspAsnValPheAspAspPheThrSerTyrAsnAl 201  
DB 384 GGACGGCACAGATGCCCTATTGATTTCTCTTCCTGATGAAAGCATACTCTGA 443  
QY 201 aHisLysLysGlnGluArg-----LysSerArgSerIleAlaAsp-- 215  
DB 444 AGAAGAAAGTAGGGCTGTTCGCCCAATAAAGCAAGCAAGCAATCAAGGACAGATGG 503  
QY 216 ----ValArgAsnGluGlnAsnIleGlnGlyAsnHisThrGluLeuGlnGluLysSe 234  
DB 216 ----ValArgAsnGluGlnAsnIleGlnGlyAsnHisThrGluLeuGlnGluLysSe 234

Db 504 CCAGTTAAACAA-GAAAAAGGAAAGAACAGCAGCCTCTCGACCAATGGCCCTC 562  
QY 234 rSerAsnGluAlaThrSerLysGluSerProAlaProLeuHisArgArgMetHi 254  
Db 563 CAGGACCCCGAGGACCTCCAGGACCCCGAGGAAATTCAGGGATTCTCTGGAA 622  
QY 254 sSerArgHisArg--HisLeuLeuValArgLysAlaArgSerGluAspSer----- 270  
Db 623 TTCAGGAACAACATGTTATGGGACCACTGGTCTCCAGGTCTCTCTGTCTCAAGGAC 682  
QY 271 -----ArgProAl 273  
Db 683 CCCTGGCTCCAGGACCTTCTGGTCTGCTGCTATAAAGCTGGAACCTCAGAAAAACAGC 742  
QY 273 a-AlaHisPheHisLeuSerSerArgArgHisGlnGlySer---MetGlyTyrHisG 292  
Db 743 CAGCTGTGGTGCATCTACAGGGC-----CAAGGGTCAGCAATTCAGTCAAGA 790  
QY 292 LyAspMetTyrIleGly-----AsnAspAsnGluArg-----AsnSerTyrGlnG 307  
Db 791 ATGATCTTCAGGTGGAGTGTCTCAATGACTGTGCTCGCATCATATGAACCCCAAGGTGT 850  
QY 307 LyHisPheGlnThrArgAspGlyValLeuThrValThrAsnThrGlyLeuTyrValT 327  
Db 851 TTAAGCTACATCCCGCAGCGGGAGCTGGAGTACTGGTGGACGCGCACTTCTCATCT 910  
QY 327 YrAlaGln-----IleCysTyrAsnAsnSerHisAspGlnAsnGlyPheIleValPheG 345  
Db 911 ATAGTCAGGTAGAAGTATATCATCACTTCACTGACTTTCAGCTATGAGGTGGTGG 970  
QY 345 LngLyAspThrProPheLeuGlnCysLeuAsnThrValProThrAsnMetProHisLysV 365  
Db 971 TGGATGAGAGCCCTTCCTGCGAGTCACAGCAGCATCGAGCGGCAAGCAACAC--T 1027  
QY 365 alHisThrCysHisThrSerGlyLeuIleHisLeuGluArgAsnGluArgIleHisLeuL 385  
Db 1028 ACAACACTTGTCTATACCGCAGCGCTGCTGCTCCTCAAGCCCGCAGAGAAGATCGCGTCA 1087  
QY 385 YAspIleHisAsnAspArgAsnAlaValLeuArgGluGlyAsnAsnArgSerTyrPheG 405  
Db 1088 AGATGTGTGACCTGACATCTCC-----ATCAACATGAGCAAGCAGCACACGCTTCTTTG 1141  
QY 405 LyIlePheLysVal 409  
Db 1142 GGGCCATCAGGCTG 1155  
RESULT 14  
ABL51009  
ID ABL51009 standard; cDNA; 1574 BP.  
XX ABL51009;  
AC ABL51009;  
DT 26-JUN-2002 (first entry)  
XX Human EDAL-II encoding cDNA SEQ ID NO:1.  
DE Human EDAL-II encoding cDNA SEQ ID NO:1.  
KW Human; mouse; EDAL-II; X-linked hypohidrotic ectodermal dysplasia; XLHED;  
KW DL receptor; DL gene; downless gene; ectodermal dysplasia;  
KW hair follicle growth; tooth growth; endocrine sweat gland development;  
KW mammary epithelial tissue growth; breast epithelium; gene; ss.  
OS Homo sapiens.  
XX US6355782-B1.  
PN 12-MAR-2002.  
XX 29-JUN-1999; 99US-00342681.  
XX 09-JUL-1998; 98US-0092279P.  
PR 15-DEC-1998; 98US-0112366P.  
XX (BAYU ) BAYLOR COLLEGE MEDICINE.



PA (UYOR-) UNIV OREGON HEALTH SCI.  
 XX Zonana J, Ferguson BM, Headon D, Overbeek P;  
 XX WPI: 2002-360478/39.  
 DR P-PSDB; ABB09082.  
 XX Novel purified hypohidrotic ectodermal dysplasias protein, useful for  
 PT promoting the development of hair follicles and tooth growth.  
 XX Example 1; Col 63-66; 86pp; English.  
 XX The present invention describes a purified hypohidrotic ectodermal  
 CC dysplasias protein (I) that promotes the development of hair follicles.  
 CC (I) is a promoter of hair follicle growth and tooth growth. (I) is useful  
 CC for the development of hair follicles. (I) is also useful for commercial  
 CC and clinical diagnostic testing for ectodermal dysplasia (ED). (I) is  
 CC also useful as a stimulant for hair and tooth growth, for stimulating  
 CC endocrine sweat gland development in individuals from whom the normal  
 CC sweating mechanism is compromised by disease or surgery, for stimulating  
 CC the growth of mammary epithelial tissue, either for reconstructing or  
 CC cosmetic purposes, and for promoting or maintaining differentiation of  
 CC breast epithelium. (I) is also useful for functional analysis, antibody  
 CC production and patient therapy. ABL51009 to ABL51122 and ABB09082 to  
 CC ABB09090 represent sequences used in the exemplification of the present  
 CC invention  
 XX SQ Sequence 1574 BP; 333 A; 497 C; 445 G; 299 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 0.0186 Length: 1574  
 Score: 138.00 Matches: 96  
 Percent Similarity: 35.92% Conservative: 57  
 Best Local Similarity: 22.54% Mismatches: 147  
 Query Match: 6.38% Indels: 127  
 DB: 6 Gaps: 18

US-09-813-329-6 (1-409) x ABL51009 (1-1574)

QY 35 ArgGlnLeuIlePro-LeuValLeuGlyPheIleGlyLeuValValAlaIleLe 54  
 DB 349 AGGNAACAGCTGCTTCTCCCTGGGTTCTTGGCTCTCGTGGCCCTCCACCTGCT 408  
 QY 54 uAlaLeuThrIleTrpGlnThrThrArgValSerHisLeuAspLysGluLeuLysSerLe 74  
 DB 409 GACGTTGTGCTGTAC-----CTAGAGTTGCGCTCGA 441  
 QY 74 uLysArgValValAspAsnLeuGlnArgLeuGlyIleAsnTyrLeuAspGluPheAs 94  
 DB 442 GTTGGCGGGGAACGTGGAGCGGAGTCCCGCTTGGC----- 478  
 QY 94 pGluPheGlnLysGluTyrGluAsnAlaLeuIleAspTyrProLysLysValAspGlyLe 114  
 DB 478 ----- 478  
 QY 114 uThrAspGluAspAspAspGlyAspGly-----LeuAs 127  
 DB 479 -----GGCTCGGCACCCCTGGCACCTCTGGCACCCCTAAG 513  
 QY 127 pSerIleAlaAsp---AspGluAspAspValSerTyrSerSerValAspValG 146  
 DB 514 CAGCTCGGTGGCTCGACCCCTGACGCCCATCACC----- 550  
 QY 146 yAlaAspTyrGluAspTyrThrAspMetLeuAsnLysLeuAsnAlaHisThrGly-- 165  
 DB 551 -----AGTCACTTGGGCA 564  
 QY 166 -ThrThrProThrSerGluThrThrAlaGluGlyGluThrAspSerAlaSerSe 185  
 DB 565 GCCGTCACTAAGCAGCAGCCATTGGACCGGGAGAGCCGCACTCCACTCTGACTCCA 624  
 QY 185 rAlaSerAsnAsp-----AspAsnValPheAspPheThrSerTyrAsnAl 201  
 DB 185 rAlaSerAsnAsp-----AspAsnValPheAspPheThrSerTyrAsnAl 201

DB 625 GGACGGGCACCATGATGGCCCTATTGAATTTCTTCTCCCTGATGAAAGCCACTACTCTGA 684  
 QY 201 aHisLysLysLysGlnGluArg-----LysSerArgSerIleAlaAsp-- 215  
 DB 685 AGAAGAAGTAGGCGTGTTCGCCCAATAAAGAAAGCAAAAGCAATGAGGAGCAGATGG 744  
 QY 216 ---ValArgAsnGluGlnAsnIleGlnGlyAsnHisThrGluLeuGlnGluLysSe 234  
 DB 745 CCCAGTTAAAAACAA-GAAAAAGGAAAGAAAGAGGACCTCTCTGGACCCCAATGGCCCTC 803  
 QY 234 rSerAsnGluAlaThrSerLysGluSerProAlaProLeuHisHisArgArgMetHi 254  
 DB 804 CAGACCCCCCAGGACCTCCAGGACCCCGAGGAAATTCAGGGATTCCTGGAA 863  
 QY 254 sSerArgHisArg---HisLeuLeuValArgLysAlaArgSerGluAspSer----- 270  
 DB 864 TTCAGGAACACTGTTATTGGGACCACCTGGTCTCTCCAGGTCCTCTCGTCTCTCAAGGAC 923  
 QY 271 -----ArgProAl 273  
 DB 924 CCCCTGGCTCCAGGACCTTCTGGTGTCTGCTGATAAAGCTGGAACTCGAGAAACACGAC 983  
 QY 273 a-AlaHisPheHisLeuSerSerArgArgArgHisGlnGlySer---MetGlyTyrHisG 292  
 DB 984 CAGCTGTGGTGCATCTACAGGC-----CAAGGCTCAGCAATTCAGTCAAGA 1031  
 QY 292 LyAspMetTyrIleGly-----AsnAspAsnGluArg-----AsnSerTyrGlnG 307  
 DB 1032 ATGATCTTTCAGGTGGAGTGTCTCAATGCTGCTCGATCTCATCACTATGAACCCCAAGGTGT 1091  
 QY 307 LyHisPheGlnThrArgAspGlyValLeuThrValThrAsnThrGlyLeuTyrTyrValT 327  
 DB 1092 TTAAGCTACATCCCGCAGCGGGAGCTGGAGGTACTGGTGGACGCGCACCTTCTCATCT 1151  
 QY 327 YrAlaGln-----IleCysTyrAsnAsnSerHisAspGlnAsnGlyPheIleValPheG 345  
 DB 1152 ATAGTCAGGTAGAAGTATACATCACTTCACTGACTTTGCCAGCTATGAGGTGGTGG 1211  
 QY 345 LngLyAspThrProPheLeuGlnCysLeuAsnThrValProThrAsnMetProHisLysV 365  
 DB 1212 TGGATGAGAGCCCTTCTCGAGTGCACAGCAGCATCGAGACGGCGGCAAGACCAAC---T 1268  
 QY 365 alHisThrCysHisThrSerGlyLeuIleHisLeuGluArgAsnGluArgIleHisLeuL 385  
 DB 1269 ACAACACTTGTATACCGCAGGCGTCTCCCTCTCAAGGCCCGCAGCAAGATCGCCGTCA 1328  
 QY 385 yAspIleHisAsnAspAsnAlaValLeuArgGluGlyAsnAsnArgSerTyrPheG 405  
 DB 1329 AGATGGTGCACGCTGACATCTCC-----ATCAACATGAGCAAGCACACCCACGTTCTTTG 1382  
 QY 405 LyIlePheLysVal 409  
 DB 1383 GGGCCATCAGGCTG 1396

RESULT 15  
 ACD07895  
 ID ACD07895 standard; cDNA; 1574 BP.  
 XX ACD07895;  
 AC ACD07895;  
 XX 15-AUG-2003 (first entry)  
 DT 15-AUG-2003 (first entry)  
 XX cDNA encoding human EDAL-II.  
 DE Human; 88; gene; EDAL-II; hair follicle development; tooth development;  
 KW sweat gland development; ectodermal disease; alopecia; skin trauma; burn;  
 KW x-linked hypohidrotic ectodermal dysplasia; XLHED; breast cancer; HED;  
 KW autosomal recessive hypohidrotic ectodermal dysplasia; hirsutism;  
 XX ectopic teeth.  
 OS Homo sapiens.  
 XX US2003023991-A1.  
 PN

XX PD 30-JAN-2003.  
 XX PF 04-DEC-2000; 2000US-00729658.  
 XX PR 09-JUL-1998; 98US-0092279P.  
 PR 15-DEC-1998; 98US-0112366P.  
 PR 29-JUN-1999; 99US-00342681.  
 XX (ZONA/) ZONANA J.  
 PA (FERG/) FERGUSON B M.  
 PA (HEAD/) HEADON D.  
 PA (OVER/) OVERBEEK P.  
 XX Zonana J, Ferguson BM, Headon D, Overbeek P;  
 XX WPI; 2003-456312/43.  
 DR P-PSDB; ABO01928.  
 XX  
 PT Increasing or decreasing hair follicle development, tooth development, or  
 PT sweat gland development, in a tissue, involves modulating ectodermal  
 PT dysplasia protein isoform II activity in the tissue.  
 XX Claim 13; Page 46-47; 102pp; English.  
 XX  
 CC The invention relates to a method of increasing or decreasing one or more  
 CC of hair follicle development, tooth development, or sweat gland  
 CC development, in a tissue, involves altering ectodermal dysplasia (EDAL)  
 CC isoform II (EDAL-II) activity in the tissue. The method is useful for  
 CC increasing or decreasing hair follicle development, tooth development or  
 CC ectodermal development, in a tissue of a patient suffering from an  
 CC ectodermal disease such as X-linked hypohidrotic ectodermal dysplasia  
 CC (XLHED), autosomal recessive hypohidrotic ectodermal dysplasia (HED) or  
 CC alopecia. The method is useful for decreasing hair follicle development  
 CC in a subject suffering from hirsutism. The method is also useful for  
 CC treating breast cancer and ectopic teeth, burns and trauma of skin due to  
 CC surgery. The present sequence represents cDNA encoding human EDAL-II  
 XX  
 SQ Sequence 1574 BP; 333 A; 497 C; 445 G; 299 T; 0 U; 0 Other;  
 Alignment Scores:  
 Pred. No.: 0.0186 Length: 1574  
 Score: 138.00 Matches: 96  
 Percent Similarity: 35.92% Conservative: 57  
 Best Local Similarity: 22.54% Mismatches: 147  
 Query Match: 6.38% Indels: 127  
 DB: 8 Gaps: 18  
 US-09-813-329-6 (1-409) x ACD07895 (1-1574)  
 QY 35 ArgGlnLeuIlePro-LeuValLeuGlyPheIleGlyLeuValValAlaIleLe 54  
 DB 349 AGGGAACAGCTGCTCTTCCTGGGTTCTTGGGCTCTCGCTGGCCCTCCACCTGCT 408  
 QY 54 uAlaLeuThrIleTrpGlnThrThrArgValSerHisLeuAspLysGlyLeuYssSerLe 74  
 DB 409 GACCTTGCTGCTAC-----CTAGAGTTGGCTCGGA 441  
 QY 74 uLysArgValValAspAsnLeuGlnArgLeuGlyIleAsnTyrLeuAspGluPheAs 94  
 DB 442 GTTCGGGGGAACGTGGAGCCGAGTCCCGCTTGGC----- 478  
 QY 94 pGluPheGlnLysGluTyrGluAsnAlaLeuIleAspTyrProLysLysValAspGlyLe 114  
 DB 478 ----- 478  
 QY 114 uThrAspGluAspAspAspGlyAspGly-----LeuAs 127  
 DB 479 -----GGCTCGGCACCCCTGGCACCTCTGGCACCCCTAAG 513  
 QY 127 pSerIleAlaAsp---AspGluAspAspAspValSerTyrSerValAspAspValG 146  
 DB 514 CAGCTCGGTGGCTCGACCTCGACCCCTGACAGCCCATCACC----- 550

QY 146 yAlaAspTyrGluAspTyrThrAspMetLeuAsnLysLeuAsnAsnAlaHisThrGly-- 165  
 DB 551 -----AGTCACCTTGGCA 564  
 QY 166 -ThrThrProThrSerGluThrThrAlaGluGlyGluGlyGluThrAspSerAlaSerSe 185  
 DB 565 GCGGTACCTTAAGCAGCAGCATTTGGAAACCGGGAAGCGCGACTCCACTCTGACTCCCA 624  
 QY 185 rAlaSerAsnAsp-----AspAsnValPheAspAspPheThrSerTyrAsnAl 201  
 DB 625 GGACGGGACACCATGCTTATTTCTTCTTCTTCTGATGAAAGCCATCTACTCTGA 684  
 QY 201 aHisLysLysLysGlnGluArg-----LysSerArgSerIleAlaAsp-- 215  
 DB 685 AGAAGAAAGTAGCGTGTTCGCGCAATAAAGAAAGCAATGAAGGAGCAGATGG 744  
 QY 216 ---ValArgAsnGluGlnAsnIleGlnGlyAsnHisThrGluLeuGlnGluLysSe 234  
 DB 745 CCAGTTTAAACAA-GAAAGGGGAAAGACGAGGACCTCTTGACCCCAATGGCCCTC 803  
 QY 234 rSerAsnGluAlaThrSerLysGluSerProAlaProLeuHisHisArgArgMethi 254  
 DB 804 CAGGACCCCGACGACCTCCAGGACCCCGAGGAATTCAGGGATTCTCTGGAA 863  
 QY 254 sSerArgHisArg---HisLeuLeuValArgLysAlaArgSerGluAspSer-- 270  
 DB 864 TTCAGGAACAACCTGTATATGGACCACTGGTCTCCAGGTCCTCTCTGCTCAAGGAC 923  
 QY 271 -----ArgProAl 273  
 DB 924 CCCTGGCCCTCCAGGACCTTCTGTGTGTGTATAGCTGGAACCTCGAACTCGAAAAACAC 983  
 QY 273 a-AlaHisPheHisLeuSerSerA-ArgArgHisGlnGlySer---MetGlyTyrHisG 292  
 DB 984 CAGCTGTGTGTCTTACAGGGC-----CAAGGGTCAGCAATTCAAGTCAAGA 1031  
 QY 292 LyAspMetTyrIleGly-----AsnAspAsnGluArg-----AsnSerTyrGlnG 307  
 DB 1032 ATGATCTTTTCAGGTGGAGTGCTCAATGACTGTGCTCGCATCATGAACCCCAAGGTGT 1091  
 QY 307 LyHisPheGlnThrArgAspGlyValLeuThrValThrAsnThrGlyLeuTyrValT 327  
 DB 1092 TTAAGCTACATCCCGCAGCGGGAGCTGGAGTACTGTGTGACCGCACCTTACTTCACT 1151  
 QY 327 YrAlaGln-----IleCysTyrAsnAsnSerHisAspGlnAsnGlyPheIleValPheG 345  
 DB 1152 ATAGTCAGGTAGAAGTATATACATCAACTTCACTGACTTTGCCAGCTATAGGTGGTGG 1211  
 QY 345 LngLysAspThrProPheLeuGlnCysLeuAsnThrValProThrAsnMetProHisLysV 365  
 DB 1212 TGGATGAGAGCCCTTCTGTGAGTCACACGAGCATCGAGACGGGGAAGACCAAC---T 1268  
 QY 365 alHisThrCysHisThrSerGlyLeuIleHisLeuGluArgAsnGluArgIleHisLeuL 385  
 DB 1269 ACAACACTTGCTATACCGAGGGGTCTCCCTCAAGGCCCGGCGAGAGATCGCGGCA 1328  
 QY 385 ysAspIleHisAsnAspAsnAlaValLeuArgGluGlyAsnAsnArgSerTyrPheG 405  
 DB 1329 AGATGGTGCAGCTGACATCTCC-----ATCAACATGAGCAAGCACACCGTCTTTTG 1382  
 QY 405 lyIlePheLysVal 409  
 DB 1383 GGGCCATCAGGCTG 1396

Search completed: September 15, 2004, 08:23:37  
 Job time : 657 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: September 15, 2004, 06:59:38 ; Search time 103 Seconds  
(without alignments)  
2203.640 Million cell updates/sec

Title: US-09-813-329-6

Perfect score: 2162

Sequence: 1 MTATLKPFFITPTTSANDGF.....DRNAVLREGNNRSYFGIFKV 409

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+P2n.model -DEV=xlp  
-Q/cgn2\_1/USPRO-spool\_p/US09813329/runat\_13092004\_102753\_21701/app\_query.fasta\_1.583  
-DB=Issued Patents NA -QMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=pco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09813329 -CGCN 1 1 128 @runat\_13092004\_102753\_21701 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGQUERY -NEG\_SCORES=0 -WAIT -DSPELOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:\*

1: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq:\*  
2: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq:\*  
3: /cgn2\_6/ptodata/2/ina/6A\_COMB.seq:\*  
4: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq:\*  
5: /cgn2\_6/ptodata/2/ina/PCTUS\_COMB.seq:\*  
6: /cgn2\_6/ptodata/2/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	138	6.4	1176	4	US-09-342-681C-14
2	138	6.4	1574	4	US-09-342-681C-1
3	137	6.3	3034	4	US-09-976-594-636
4	133	6.2	3489	2	US-08-728-323A-1
5	133	6.2	3489	2	US-09-298-568-1
6	133	6.2	3489	4	US-09-410-399-1
7	133	6.2	32207	2	US-08-770-379-20
8	133	6.2	32207	3	US-08-757-669A-20
9	133	6.2	32207	4	US-09-230-371A-20
10	129	6.0	4599	1	US-08-431-080-27
11	129	6.0	4599	2	US-08-938-534-27
12	129	6.0	4599	4	US-09-345-294-27

13	128.5	5.9	4522	4	US-08-956-171E-39	Sequence 39, Appl
14	123	5.7	1422	4	US-09-485-077A-15	Sequence 15, Appl
15	121.5	5.6	1424	4	US-09-206-576-1	Sequence 1, Appl
16	120	5.6	4507	2	US-08-588-459A-3	Sequence 3, Appl
17	120	5.6	4507	2	US-08-487-826B-3	Sequence 3, Appl
18	120	5.6	4507	4	US-09-210-288-3	Sequence 3, Appl
19	119	5.5	2854	4	US-08-971-089-3	Sequence 3, Appl
20	118	5.5	1493	6	5340934-5	Patent No. 5340934
21	117.5	5.4	866	4	US-09-589-287B-37	Sequence 37, Appl
22	117.5	5.4	866	4	US-09-588-947A-37	Sequence 37, Appl
23	117.5	5.4	866	4	US-09-589-286A-37	Sequence 37, Appl
24	116	5.4	1236	3	US-09-105-343A-1	Sequence 1, Appl
25	116	5.4	2017	1	US-07-667-276A-3	Sequence 3, Appl
26	115.5	5.3	1478	4	US-09-216-393B-7	Sequence 7, Appl
27	115	5.3	30549	4	US-09-134-001C-322	Sequence 322, App
28	114.5	5.3	564	4	US-09-396-937-7	Sequence 7, Appl
29	114.5	5.3	4403765	3	US-09-103-840A-2	Sequence 2, Appl
30	114.5	5.3	4411529	3	US-09-103-840A-1	Sequence 1, Appl
31	114	5.3	1512	4	US-09-252-991A-8290	Sequence 8290, Ap
32	114	5.3	2637	4	US-09-540-236-1481	Sequence 1481, Ap
33	113	5.2	956	1	US-08-431-080-25	Sequence 25, Appl
34	113	5.2	956	1	US-08-431-080-32	Sequence 32, Appl
35	113	5.2	956	2	US-08-938-534-25	Sequence 25, Appl
36	113	5.2	956	2	US-08-938-534-32	Sequence 32, Appl
37	113	5.2	956	2	US-09-345-294-25	Sequence 25, Appl
38	113	5.2	956	4	US-09-345-294-32	Sequence 32, Appl
39	113	5.2	49617	4	US-09-596-002-28	Sequence 28, Appl
40	112.5	5.2	1952	4	US-09-533-029-89	Sequence 89, Appl
41	112.5	5.2	2384	1	US-07-814-964-10	Sequence 10, Appl
42	112.5	5.2	2384	1	US-08-258-442-10	Sequence 10, Appl
43	112.5	5.2	2384	1	US-08-328-809-5	Sequence 5, Appl
44	112.5	5.2	2384	4	US-08-866-840-5	Sequence 5, Appl
45	112.5	5.2	2384	5	PCT-US92-11107-10	Sequence 10, Appl

ALIGNMENTS

RESULT 1

US-09-342-681C-14  
; Sequence 14, Application US/09342681C  
; Patent No. 6355782  
; GENERAL INFORMATION:  
; APPLICANT: Zonana et al.  
; TITLE OF INVENTION: Hypohydrotic ectodermal dysplasia genes and proteins  
; FILE REFERENCE: 52978  
; CURRENT APPLICATION NUMBER: US/09/342,681C  
; CURRENT FILING DATE: 1999-06-29  
; PRIOR APPLICATION NUMBER: 60/092,279  
; PRIOR FILING DATE: 1998-07-09  
; PRIOR APPLICATION NUMBER: 60/112,366  
; PRIOR FILING DATE: 1998-12-15  
; NUMBER OF SEQ ID NOS: 123  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 14  
; LENGTH: 1176  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(1176)  
US-09-342-681C-14

Alignment Scores:  
Pred. No.: 2.12e-05  
Score: 138.00  
Percent Similarity: 35.92%  
Best Local Similarity: 22.54%  
Query Match: 6.38%  
DB: 4  
Length: 1176  
Matches: 96  
Conservative: 57  
Mismatches: 147  
Indels: 127  
Gaps: 18

US-09-813-329-6 (1-409) x US-09-342-681C-14 (1-1176)

Qy 35 ArgGlnLeuIlePro-LeuValLeuGlyPheIleGlyLeuGlyValValAlaIleLe 54



```
Qy 146 yAlaaspTyrGluaspTyrThrAspMetLeuAsnLysLeuAsnAlaHisThrGly-- 165
Db 551 -----AGTCACTTGGGCA 564
Qy 166 -ThrThrProThrSerGluThrThrAlaGluGlyGluThrAspSerAlaSerSe 185
Db 565 GCCGTCACTTAAGCAGCAGCATTTGGAAACCGGGAGAGCCGACCTCCACTCTGACTCCCA 624
Qy 185 rAlaSerAsnAsp-----AspAsnValPheAspAspPheThrSerTyrAsnAl 201
Db 625 GGACGGGCACACAGATGGCCCTATTGAATTTCTTCCCTGATGAAAGCCATACTCTGA 684
Qy 201 aHisLysLysLysGlnGluArg-----LysSerArgSerIleAlaAsp-- 215
Db 685 AGAAGAAAGTAGGCGGTGTTCCGCCCAATAAAGCAAGCAATGATGAAGCAGCAGATGG 744
Qy 216 ---ValArgAsnGluGlnAsnIleGlnGlyAsnHisThrGluLeuGlnGluLysSe 234
Db 745 CCCAGTTTAAAAACAA-GAAAGAGGGAAGAAAGAGCAGGACCTCCTGGACCCCAATGGCCCTC 803
Qy 234 rSerAsnGluAlaThrSerLysGluSerProAlaProLeuHisHisArgArgMetHi 254
Db 804 CAGGACCCCGAGGACCTCCAGGACCCCGAGGACCCCGAGGATTCAGGGATTCCTGGAA 863
Qy 254 sSerArgHisArg---HisLeuLeuValArgLysAlaArgSerGluAspSer----- 270
Db 864 TTCAGGAACAACATGTTATATGGGACCACTGGTCTCCAGGTCTCTGCTCAAGGAC 923
Qy 271 -----ArgProAl 273
Db 924 CCCTCGCCCTCCAGGACCTTCTGTGCTGCTGATAAAGCTGGAACCTCGAAGAACCCAGC 983
Qy 273 a-AlaHisPheHisLeuSerSerArgArgHisGlnGlySer---MetGlyTyrHisG 292
Db 984 CAGCTGTGGTGCATCTACAGGC-----CAGGGTCAAGCAATTCAGTCAAGA 1031
Qy 292 lyAspMetTyrIleGly-----AsnAspAsnGluArg-----AsnSerTyrGlnG 307
Db 1032 ATGATCTTTCAGGTGGAGTGCTCAATGACTGGTCTCGCATCACTATGAACCCCAAGGTGT 1091
Qy 307 lyHisPheGlnThrArgAspGlyValLeuThrValThrAsnThrGlyLeuTyrTyrValT 327
Db 1092 TTAAGCTATACATCCCGCAGCGGGAGCTGGAGGTACTGGTGGAGCGGACCTTACTTCACT 1151
Qy 327 yAlaGln-----IleCysTyrAsnAsnSerHisAspGlnAsnGlyPheIleValPheG 345
Db 1152 ATATCTCAGGTAGAAGTATACATCACTTCACTGACTTTGCCAGCTATGAGGTGGTG 1211
Qy 345 lnglyAspThrProPheLeuGlnCysLeuAsnThrValProThrAsnMetProHisLysV 365
Db 1212 TGGATGAGAGCCCTTCTCGCAGTGCACAGCAGCATCGAGACGGGCAAGACCAAC---T 1268
Qy 365 alHisThrCysHisThrSerGlyLeuIleHisLeuGluArgAsnGluArgIleHisLeuL 385
Db 1269 ACAACACTGTATATACCGCAGCGCTGCTCCCTCAAGCGCCCGCAGAGATCCGCCGTCA 1328
Qy 385 yAspIleHisAsnAspArgAsnAlaValLeuArgGluGlyAsnAsnArgSerTyrPheG 405
Db 1329 AGATGGTGACGCTGACATCTCC-----ATCAACATGAGCAGCACCACCGTTCITTG 1382
Qy 405 lyIlePheLysVal 409
Db 1383 GGGCCATCAGGCTG 1396
```

## RESULT 3

```
US-09-976-594-636
; Sequence 636, Application US/09976594
; Patent No. 6673549
; GENERAL INFORMATION:
; APPLICANT: Furness, Michael
; APPLICANT: Buchbinder, Jenny
; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
; FILE REFERENCE: PA-0041 US
```

```
; CURRENT APPLICATION NUMBER: US/09/976,594
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 60/240,409
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 1143
; SOFTWARE: PERL Program
; SEQ ID NO 636
; LENGTH: 3034
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6673549 238877.11
US-09-976-594-636
```

## Alignment Scores:

```
Pred. No.: 0.000119 Length: 3034
Score: 137.00 Matches: 68
Percent Similarity: 36.56% Conservative: 49
Best Local Similarity: 21.25% Mismatches: 107
Query Match: 6.34% Indels: 96
DB: 4 Gaps: 12
```

US-09-813-329-6 (1-409) x US-09-976-594-636 (1-3034)

```
Qy 20 PheProAlaLysAlaThrSerThrAlaThrAlaGlnArgThrArgGlnLeuIlePro 39
Db 332 TTTCCTAGAGAACACCAGCTGTAATCGATGTCGTCGGGAGAAACAACATGAGGCCAAGATG 391
Qy 40 LeuValLeuGlyPheIleGlyLeuGlyLeuValValAla-----IleLeu 54
Db 392 ATGAAGCTGGGCGCACTGAATAGGAAGACACCTTCAGAAAAGAGCCGAGGCTATTT 451
Qy 55 AlaLeuThrIleTrpGlnThrThrArgValSerHisLeuAspLysGluLeuLysSerLeu 74
Db 452 AGTGCTAATGACTGGCAATGTAACACTTGCAGCAATGTGAATTTGGGCCAGAGATCAGAG 511
Qy 75 LysArgValValAla-----AsnLeuGlnGlnArgLeuGly----- 86
Db 512 TGTAATATGTGTAATACTCCAAAGTATGCTAAATAGAAAGAACAGACAGATATGGTGGT 571
Qy 87 -----IleAsnTyrLeuAsp-----GluPhe 93
Db 572 GGTTTTAAATCAAGAGAGAAAATGTTGAATATATAGAAAGAGAGAAGAAATCTGATGGTGAATAT 631
Qy 94 AspGluPheGlnLysGluTyrGluAsnAlaLeuIleAspTyrProLysLysValAspGly 113
Db 632 GATGAGTTTGGACGTAAAAAGAAAAA-----TACAGAGGGAAGCAGATTGGT 679
Qy 114 -----LeuThrAspGluGluAspAspAspGlyAspGlyLeuAspSerIleAla 130
Db 680 CTTGCATCTATATAAGGAGCTTGAAGATAAGAAATCAGAGGGAGAGAA----- 730
Qy 131 AspAspGluAspAspAspValSerTyrSerValAspAspValGlyAlaAspTyrGlu 150
Db 731 GAGGATGAGGATCAAGATCTTTCTAAATATAAGTAGAT----- 769
Qy 151 AspTyrThrAspMetLeuAsnLysLeuAsnAsnAlaHisThrGlyThrThrProThrSer 170
Db 769 ----- 769
Qy 171 GluThrThrAlaGluGlyGluGlyGluThrAspSerAlaSerSerAlaSerAsnAspAsp 190
Db 770 -----GAGGATGAGGATGAAGATGAGCTGATCTCTCAAAATATAT- 811
Qy 191 AsnValPheAspAspPheThrSerTyrAsnAlaHisLysLysLysGlnGluArgLysSer 210
Db 812 -----CTTCATGCCAGTGAAGAGAGAGATAGTAATAAAAAAGAAATCTAATAGACGAAGT 865
Qy 211 ArgSerIleAlaAspValArgAsnGluGluGlnAsnIleGlnGlyAsnHisThrGluLeu 230
Db 866 CGTCA----- 871
```

QY 231 GlnGluYsSerSerAsnGluAlaThrSerLysGluSerProAlaProLeuHisHisArg 250  
 Db 872 AAGTCTCGATCTTCAATTCAGATCTTCATCAGCTCATCTCTCCCTCAAGTTCAGG 931  
 QY 251 ArgArgMetHisSerArgHisArgHisLeuValArgLysAlaArgSer----- 267  
 Db 932 TCTAGGTCAGGTCCTCGTTCAGGAAGTCTTCCAGTTCGCAGTCAAGATCTCGTTCCAGT 991  
 QY 268 -----GluAspSerArgProAlaHisPheHisLeuSerSerArgArgHisGln 285  
 Db 992 TCCAGAGAACGTTCCAGATCTCGTGGGTCGAAATCAAGATCCAGTCCAGTCCACAGG 1051  
 QY 286 GlySerMetGlyTyHisGlyAspMetTyTrileGlyAsnAspAsn-----GluArgAsn 303  
 Db 1052 GGCCTCTTCTCCACGAAAAAGATCTTATTCAGTTTCATCATCTTCTCTGAGAGGAAC 1111

## RESULT 4

US-08-728-323A-1  
 ; Sequence 1, Application US/08728323A  
 ; Patent No. 5948676  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Chang, Yuan  
 ; APPLICANT: Bohenzky, Roy A.  
 ; APPLICANT: Russo, James J.  
 ; APPLICANT: Edelman, Isidore S.  
 ; APPLICANT: Moore, Patrick S.  
 ; TITLE OF INVENTION: Immediate Early Protein From Kaposi's  
 ; TITLE OF INVENTION: Sarcoma-Associated Herpesvirus, DNA  
 ; TITLE OF INVENTION: Encoding Same And Uses Thereof  
 ; NUMBER OF SEQUENCES: 21  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Cooper & Dunham LLP  
 ; STREET: 1185 Avenue of the Americas  
 ; CITY: New York  
 ; STATE: New York  
 ; COUNTRY: U.S.A.  
 ; ZIP: 10036  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patentin Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/728,323A  
 ; FILING DATE:  
 ; CLASSIFICATION: 435  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: White, John P.  
 ; REGISTRATION NUMBER: 28,678  
 ; REFERENCE/DOCKET NUMBER: 0575/52268/JPW/MSC/SKS  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 212-278-0400  
 ; TELEFAX: 212-391-0525  
 ; INFORMATION FOR SEQ ID NO: 1:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 3489 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: DNA (genomic)  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: 1..3489  
 US-08-728-323A-1

Alignment Scores:  
 Pred. No.: 0.000405  
 Score: 133.00  
 Percent Similarity: 35.47%  
 Best Local Similarity: 25.21%  
 Query Match: 6.15%  
 DB: 2

US-09-813-329-6 (1-409) x US-08-728-323A-1 (1-3489)  
 QY 91 AspGluPheAspGluPheGlnLysGluTyArgGluAsnAlaLeuLeAspTyProLysLys 110  
 Db 970 GATGACAATGACAATAGGATGATGAGGAGGACGAGACACATGAGGAGGACGAGGAG 1029  
 QY 111 ValAspGlyLeuThrAspGluAspAspAspAspGlyAspGlyLeuAspSerIleAla 130  
 Db 1030 GATGACGAGGAGGATGACGAGGAGGATGACGAGGAGGATGACGAGGAGGATGACGAGGAG 1089  
 QY 131 AspAspGluAspAspValSerTySerSerValAspAspValGlyAlaAspTyArgLys 150  
 Db 1090 GATGACGAGGAGGATGACGAGGAGGATGACGAGGAGGATGACGAGGAGGATGACGAGGAG 1149  
 QY 151 AspTyThrAspMetLeuAsnLysLeuAsnAlaHisThrGlyThrThrProThrSer 170  
 Db 1150 GAT-----GACGAG 1158  
 QY 171 GluThrThrAlaGluGlyGluThrAspSerAlaSerSerAlaSerAsnAspAsp 190  
 Db 1159 GAGGAGGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGATGAC 1218  
 QY 191 AsnValPheAspAspPheThrSerTyArgAsnAlaHisLysLysLysGlnGluArgLysSer 210  
 Db 1219 -----GATGATGAGGACATGAGGAGGAGGATGACGAGGAGGAGGAGGAGGAGGAGGAG 1269  
 QY 211 ArgSerIleAlaAspValArgAsnGluGlnAsnIleGlnGlyAsnHisThr----- 228  
 Db 1270 -----GAGGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1311  
 QY 229 -----GluLeuGlnGlu-LysSer-----SerAsnGluAlaTh 239  
 Db 1312 ATCCAAAGTTCACACAGCAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1371  
 QY 239 rSerLysGluSerProAlaProLeuHis-----HisArgArgArgMethHisSe 255  
 Db 1372 CAGCAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1431  
 QY 255 rArgHisArgHisLeuValArgLysAlaArgSerGluAspSerArgProAlaAlaHis 275  
 Db 1432 CAGCAGGAG 1491  
 QY 275 sPheHisLeuSerSerArgArgArgHisGlnGlySerMetGlyTyThrHisGlyAspMetTy 295  
 Db 1492 GAGCCAC---AACAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1548  
 QY 295 rIleGlyAsnAspAsnGluArgAsnSerTyTrileGlyHis 308

## RESULT 5

US-09-298-568-1  
 ; Sequence 1, Application US/09298568  
 ; Patent No. 6322792  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Kieff, Elliott D.  
 ; APPLICANT: Ballestas, Mary E.  
 ; APPLICANT: Kaye, Kenneth M.  
 ; TITLE OF INVENTION: RHADINO VIRUS LANA ACTS IN TRANS ON A UNIT OF RHADINO  
 ; TITLE OF INVENTION: VIRUS DNA TO MEDIATE EFFICIENT EPISOME PERSISTENCE  
 ; FILE REFERENCE: 16412-10001R  
 ; CURRENT APPLICATION NUMBER: US/09/298,568  
 ; EARLIER FILING DATE: 1999-04-21  
 ; EARLIER APPLICATION NUMBER: US 60/109,422  
 ; EARLIER FILING DATE: 1998-11-19  
 ; NUMBER OF SEQ ID NOS: 3  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 1  
 ; LENGTH: 3489  
 ; TYPE: DNA  
 ; ORGANISM: Kaposi's sarcoma-associated herpesvirus  
 US-09-298-568-1



```

/ ADDRESS: Cooper & Dunham LLP
/ STREET: 1185 Avenue of the Americas
/ CITY: New York
/ STATE: New York
/ COUNTRY: U.S.A.
/ ZIP: 10036
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patentin Release #1.0, Version #1.30
/ CURRENT APPLICATION NUMBER: US/08/757,669A
/ FILING DATE:
/ CLASSIFICATION: 424
/ ATTORNEY/AGENT INFORMATION:
/ NAME: White, John P.
/ REGISTRATION NUMBER: 28,678
/ REFERENCE/POCKET NUMBER: 45185-F
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (212) 278-0400
/ TELEFAX: (212) 391-0525
/ INFORMATION FOR SEQ ID NO: 20:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 32207 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: double
/ TOPOLOGY: linear
/ MOLECULE TYPE: DNA (genomic)
/ US-08-757-669A-20

Alignment Scores:
Pred. No.: 0.0129 Length: 32207
Score: 133.00 Matches: 59
Percent Similarity: 35.47% Conservatives: 24
Best Local Similarity: 25.21% Mismatches: 108
Query Match: 6.15% Indels: 43
DB: 3 Gaps: 7

```









```
i
US-09-345-294-27
SEQUENCE DESCRIPTION: SEQ ID NO: 27:
Alignment Scores:
Pred. No.: 0.0017 Length: 4599
Score: 129.00 Matches: 73
Percent Similarity: 32.97% Conservative: 48
Best Local Similarity: 19.83% Mismatches: 126
Query Match: 5.97% Indels: 120
DB: 4 Gaps: 12

US-09-813-329-6 (1-409) x US-09-345-294-27 (1-4599)
QY 69 LysGluLeuLysSerLeuLysArgValValAspAsnLeuGlnArgLeuGlyIleAsn 88
DB 1114 AAAAAAGTAACTTATCAAGGCGAATAGACAAATGACGAGGCGC-----1161

QY 89 TyrLeuAspGluPheAspGluPheGln-----LysGluTyrGluAsnAla 103
DB 1162 -----ACAGAGTCATCCGATTACCAAGCTGTAACAGATGCGAAGAGAGTGAATAAGAA 1215

QY 104 LeuIleAspTyrProLysLysValAspGlyLeuThrAspGluGluAspAspAspGly 123
DB 1216 GAAGAAGAGAGTGAAGAAGAAGAAGATGATGACGAAGACGACGACGATGATGAT 1275

QY 124 AspGlyLeuAspSerIleAlaAspAspGluAspAspValSerTyrSerSerValAsp 143
DB 1276 GACGGAAGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1317

QY 144 AspValGlyAlaAspTyrGluAsp-----151
DB 1318 GACGAGAATATAGATTTGTCAAAGTACGAGCCCAAGAAAGAAAGAGGCGATGAAAGCT 1377

QY 152 -----TyrThrAspMetLeuAsnLysLeuAsn 160
DB 1378 TTATCTGCCATGATACAAATAGCAATACACTTATTCCTCTCTGAGAACAGTAAACAA 1437

QY 161 AsnAlaHisThrGlyThrProThrSerGluThrThrAlaGlu-----175
DB 1438 AATAAATCAGTTAAATATATCCCTTAAGAAAGAAATGAAGAGAGCAGAGAAGAA 1497

QY 176 -----Gly 176
DB 1498 GAAAAAGAGAAGAGAGCAACAAAAACAAAGAAATCAAAACAAAAAAGAAAGCGT 1557

QY 177 GluGlyGluThrAspSerAlaSerSerAlaSerAsnAspAspAsnValPheAspPhe 196
DB 1558 TCAGGCACACTACTACTACACACAGCGCTATCGTTTAAATTCAAAAAGAGGACGCGC 1617

QY 197 ThrSerTyr---AsnAlaHisLysLysLysGluArgLysSerArgSerIleAlaAsp 215
DB 1618 ATTAGTTTGGTAATGTAATGAGGCTATACAGAGATATAGGTGAAGAGTCTGGAT 1677

QY 216 ValArgAsnGluGluGlnAsn-----IleGln 224
DB 1678 TTAATAAACAAAGAGAACAAATGTAATCAAGAGATAAACTGGATTCTAAGGTGATGTTA 1737

QY 225 GlyAsnHisThrGluGluLysSerSerAsnGluAlaThrSerLysGluSerPro 244
DB 1738 GGTAAACAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1791

QY 245 AlaProLeuHisHisArgArgArgMetHisSerArgHisHisLeuLeuValArgLys 264
DB 1791 -----1791

QY 265 AlaArgSerGluAspSerArgProAlaAlaHisPheHisLeuSerSerArgArgHis 284
DB 1792 -----TATGATATTGACAGGATGCGTACTTACGCTGATTAACATGAAGATTCT 1842

QY 285 GlnGlySerMetGlyTyr-----HisGlyAspMetTyrIleGlyAsnAsp 299
DB 1843 CATGAGAAATGGTACAGATCTTGAAACCGGGGACAGCATCTTCCCATATTGGAGAA 1902
```

```

300 AsnGluArgAsn-----SerTyrGlnGly---- 307
1903 GAAGAACAAACAAATGTTTCTGAGCTACAAAATGACGACGAACTCTCATTCGATGCTAGT 1962
308 ---HisPheGlnThrArgAspGlyValLeuThrValThrAsnThrGlyLeuTyrVal 326
1963 ATACACGAAGAGGGTCTGATCTCTGAGAGATGCTGAAANATAAATTTTTCGAAAATGAA 2022
327 TyrAlaGln---IleCysTyrAsnAsnSerHisAspGlnAsnGlyPheIleValPheGln 345
2023 TACAATCAAGAAACGATATGAGGAAGATGACGAAAGAGATGAAATATATGCTGAT 2082
346 GlyAspThrProPheLeuGln 352
2083 TTTGATATGCCGTTTATGAA 2103

RESULT 13
US-08-956-171E-39
; Sequence 39, Application US/08956171E
; Patent No. 6593114
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; Gil H. Choi
; Patrick S. Dillon
; Craig A. Rosen
; Steven C. Barash
; Michael R. Fannon
; ADDRESS: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/956,171E
; FILING DATE: 20-Oct-1997
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/009,861
; FILING DATE: January 5, 1996
; APPLICATION NUMBER: 08/781,986
; FILING DATE: January 3, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mark J. Hyman
; REGISTRATION NUMBER: 46,789
; REFERENCE/DOCKET NUMBER: PB248P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (240) 314-1224
; TELEFAX: (301) 309-8439
; INFORMATION FOR SEQ ID NO: 39:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4522 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 39:
US-08-956-171E-39

Alignment Scores:
Pred. No.: 0.00187 Length: 4522
Score: 128.50 Matches: 73
Percent Similarity: 30.25% Conservative: 48
Best Local Similarity: 18.25% Mismatches: 120
Query Match: 5.94% Indels: 159
DB: 4 Gaps: 13
```

```
US-09-813-329-6 (1-409) x US-08-956-171E-39 (1-4522)
Qy 60 GlnThrThrArgValSerHisLeuAspLysGluLeuLysSerLeuLysArgValValAsp 79
Db 2194 CAACMAAGCGTGAAGAACATAAACCAGGAGCAAGATGCAATACAAACGTCGAATTGAT 2253
Qy 80 AsnLeuGlnArgLeuGlyLeuAsnTyrLeu
Db 2254 GAAATGTATGCTAAACACGCGCAAGCTATGTTGGTGATGATGTCATTAAATGATCATAGT 2313
Qy 90
Db 2314 GACTTACAGATAATAGTACAGATGCTAGTCAGCTTCATACAAATGGCATAGAGAATGAA 2373
Qy 91
Db 2374 ACTGTATCAATGATGAATAATAACMACGGCTCAATACAAATGAAGACACTAAATGACACT 2433
Qy 93 ---PheAspGluPheGlnLysGluTyrGluAsnAlaLeuIleAspTyrProLysLysVal 111
Db 2434 CATGTAGATGAAGTCCATACAAATATATGAGGAGTTAGTTTGAATCAAGTATCGACAACA 2493
Qy 112 AspGlyLeuThrAspGluGlu
Db 2494 AAACAAATGTCAGATGATGAAGTTACGGTTTCGAATGTGAAGTCTCAACATCAATCAGCA 2553
Qy 119
Db 2554 CTACACATAACGTTGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAAT 2613
Qy 130 AlaAspAspGluAspAsp
Db 2614 GCTGATTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2673
Qy 144 AspValGlyAlaAspTyrGluAspTyrThrAspMetLeuAsnLysLeuAsnAlaHis 163
Db 2674 GAT
Qy 164 ThrGlyThrThrProThrSerGluThrAlaGluGlyGluGlyGluThrAsp
Db 2701
Qy 182
Db 2755 CGCTGTAATGTCATAGACGCTCTTAACGCAAGAGTACGAGTAATCAATCTCA 2814
Qy 186 AlaSerAsnAspAspAsnValPheAspAspPheThrSerTyrAsnAlaHisLysLys 205
Db 2815 GTAAATAATATCGATAAAACCGTTGATAATGAATGAATGAATGAATGAATGAATGAATGAAT 2874
Qy 206 GlnGluArgLysSerArgSerIleAlaAspValArgAsnGluGluGlnAsn
Db 2875 GATGACCAACAAACTTAAAGTGTCAACTCTTGAAGAACGAATGATGTAATGATATCAT 2934
Qy 223
Db 2935 GTTGTGGAGATTCAAGCATGAATCAATGAAGAAAGATAACGAGAAATTTACAGAAAT 2994
Qy 234 SerSerAsnGluAlaThrSerLysGlu
Db 2995 GTGCAAAACGAGACGCTGAAAGTGACAAATAATGTCGAAGAGAAATCTATTGAAACGTA 3054
Qy 243
Db 3055 AATCCAAAGAACACAGACTGAAAGGTTTCAACTTTAATGAAAGACCACTTAAATGTTGTC 3114
Qy 249
Db 3115 ATGACGCCATCTGATAAAACCGTATGATGATGATGATGATGATGATGATGATGATGATGAT 3174
Qy 264
Db 3175 CCTGAATTAAGCCTGTACAAAGTAAGCAAGCTGTGAGTGAAAGATCCCTCGCAGTCAA 3234
Qy 276 PheHisLeuSerSerArgArgArgHisGlnGlySer
Db 3235 GCCACACCATCATCAGATCTGATTCAAGAGTCAATACAAATGATATATAAACAAAT 3294
Qy 294 MetTyrIleGlyAsnAspAsnGluArgAsnSerTyrGlnGlyHisPheGlnThrArgAsp 313
Db 3295 AATATGACATCAAAACAATGTTGAGAACAACTTATTGGTCATGCAGAAACAGAAAT 3354

RESULT 14
US-09-485-077A-15
; Sequence 15, Application US/09485077A
; Patent No. 6458590
; GENERAL INFORMATION:
; APPLICANT: Mukherjee, Anil
; APPLICANT: Kundu, Gopal
; APPLICANT: Panda, Dibyendu
; TITLE OF INVENTION: Methods and Compositions for Treatment of Restenosis
; FILE REFERENCE: NIH-05047
; CURRENT APPLICATION NUMBER: US/09/485,077A
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: PCT/US98/16569
; PRIOR FILING DATE: 1998-07-08
; PRIOR APPLICATION NUMBER: 60/054,967
; PRIOR FILING DATE: 1997-07-08
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 15
; LENGTH: 1422
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-485-077A-15

Alignment Scores:
Pred. No.: 0.00123 Length: 1422
Score: 123.00 Matches: 79
Percent Similarity: 33.14% Conservative: 38
Best Local Similarity: 22.38% Mismatches: 116
Query Match: 5.69% Indels: 120
DB: 4 Gaps: 17

US-09-813-329-6 (1-409) x US-09-485-077A-15 (1-1422)
Qy 21 ProAlaLysAlaThrSerThrAlaThrAlaGlnArgThrArgGlnLeuIleProLeu 40
Db 19 CCAGTTGCAGCTTCTCAGCCAAACCCGACCAAGAAAACCTCCTACCATGAGAGATTGCA 78
Qy 41 ValLeuGlyPheIleGlyLeuValValAlaIleLeuAlaLeuThrIleTrpGln 60
Db 79 GTGATTTGCTTTTGCCTCTAGGCATCACCTGTGCCATA-----CCAGTTAAACAG 129
Qy 61 ThrThrArgValSerHisLeuAspLysGluLeu----- 71
Db 130 GCTGATTCTGAAAGTTCTGAGGAAAGACAGCTTTACAAACAATACCCAGATGCTGTGCC 189
Qy 72 -----LysSerLeuLysArgValValAspAsnLeuGlnArgLeu 85
Db 190 ACATGGCTAAACCTGACCATCTCAGAACAG-----AATCTCTAGCCCCACAG 240
Qy 86 GlyIleAsnTyrLeuAspGluPheGlnLysGluTyrGluAsnAlaLeuIle 105
Db 241 AATGCTGTGCTCTGAAGAACCAATGACTTTTAAACAACAG----- 282
Qy 106 AspTyrProLysLys-----ValAspGlyLeuThrAspGluGlu 118
Db 283 ACCCTTCCAAAGTAAGTCCAAACGAAAGCCATGACCATGATGATGATGATGATGATGAT 342
Qy 119 AspAspAspAspGlyAspGlyLeuAspSerIleAlaAspAspGluAspAspVal--- 137
Db 343 GATGATGACCATGTGGACGCCAGGACTCCATTGACTCGAACGACTCTGATGATGATGAT 402
Qy 138 -----SerTyrSerSerValAspValGlyAlaAspTyrGluAsp----- 151
```



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: September 15, 2004, 10:23:37 ; Search time 744 Seconds  
(without alignments)  
2765.985 Million cell update/sec

Title: US-09-813-329-6  
Perfect score: 2162  
Sequence: 1 MTATLKPFTTPTTSANDGF.....DRNAVLREGNRSYFGIFKV 409

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 3304383 seqs, 2515761380 residues

Total number of hits satisfying chosen parameters: 6608766

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+\_p2n.model -DEV=xlp  
-Q=/cgn2\_1/USPTO\_spool\_p/US09813329/runat\_13092004\_102755\_21784/app\_query.fasta\_1.583  
-DB=Published Applications NA -QFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62  
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100  
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=2000000000 -USER=US09813329 @CGN 1 1 912 @runat\_13092004\_102755\_21784  
-NCPU=6 -ICPU=3 -NO\_WMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPEL=100  
-LONGLOG -DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA.\*  
1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq.\*  
3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq.\*  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq.\*  
6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq.\*  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq.\*  
10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq.\*  
11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq.\*  
12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq.\*  
13: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq.\*  
14: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq.\*  
15: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq.\*  
16: /cgn2\_6/ptodata/2/pubpna/US10C\_PUBCOMB.seq.\*  
17: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq.\*  
18: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq.\*  
19: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	ID	Description
------------	-------------	-------	--------	----	-------------

ALIGNMENTS

RESULT 1  
US-09-813-329-5  
; Sequence 5, Application US/09813329  
; Patent No. US20020012968A1  
; GENERAL INFORMATION:  
; APPLICANT: Bristol-Myers Squibb Company  
; TITLE OF INVENTION: No. US20020012968A1el Drosophila Tumor Necrosis Factor Class Mole.  
; TITLE OF INVENTION: Variants thereof  
; FILE REFERENCE: D0016.np  
; CURRENT APPLICATION NUMBER: US/09/813,329  
; CURRENT FILING DATE: 2001-03-20  
; PRIOR APPLICATION NUMBER: 60/190,816  
; PRIOR FILING DATE: 2000-03-21  
; NUMBER OF SEQ ID NOS: 65  
; SOFTWARE: Patentin version 3.0  
; SEQ ID NO 5  
; LENGTH: 2148  
; TYPE: DNA  
; ORGANISM: Drosophila melanogaster  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (634)..(1860)  
US-09-813-329-5

1	2162	100.0	2148	9	US-09-813-329-5	Sequence 5, Appli	
2	2082.5	96.3	1221	9	US-09-813-329-3	Sequence 3, Appli	
3	2045	94.6	2166	9	US-09-813-329-1	Sequence 1, Appli	
C	4	145.5	6.7	1765	17	US-10-437-963-45792	Sequence 45792, A
5	143	6.6	1269	13	US-10-424-599-46034	Sequence 46034, A	
6	141	6.5	1287	17	US-10-451-467A-447	Sequence 447, App	
7	140.5	6.5	1221	17	US-10-437-963-52112	Sequence 52112, A	
8	138.5	6.4	1057	13	US-10-424-599-81649	Sequence 81649, A	
9	138	6.4	966	13	US-10-425-114-7734	Sequence 7734, Ap	
10	138	6.4	1176	10	US-09-729-658B-14	Sequence 14, Appl	
11	138	6.4	1574	10	US-09-729-658B-1	Sequence 1, Appli	
12	138	6.4	5307	13	US-10-202-062-41	Sequence 41, Appl	
13	138	6.4	5307	15	US-10-272-411-6	Sequence 6, Appli	
14	138	6.4	5307	15	US-10-218-547-41	Sequence 41, Appl	
15	138	6.4	5307	15	US-10-272-328A-6	Sequence 6, Appli	
16	138	6.4	5307	15	US-10-310-793-45	Sequence 45, Appl	
C	17	137	6.3	1845	15	US-10-029-386-24791	Sequence 24791, A
18	135	6.2	1626	9	US-09-764-848-19	Sequence 19, Appl	
19	135	6.2	1626	15	US-10-116-016-19	Sequence 19, Appl	
20	135	6.2	1626	15	US-10-103-313-271	Sequence 271, App	
21	135	6.2	1626	15	US-10-222-020-19	Sequence 19, Appl	
22	133	6.2	3489	13	US-09-894-273-1	Sequence 1, Appli	
23	133	6.2	3489	15	US-10-294-804-1	Sequence 1, Appli	
24	132	6.1	1681	13	US-10-424-599-107179	Sequence 107179, A	
25	129.5	6.0	2091	17	US-10-437-963-87071	Sequence 87071, A	
26	129	6.0	1508	13	US-10-424-599-69258	Sequence 69258, A	
27	128.5	5.9	2151	17	US-10-437-963-60017	Sequence 60017, A	
28	128.5	5.9	3100	17	US-10-479-435-58	Sequence 58, Appl	
29	128.5	5.9	4522	8	US-08-781-986A-39	Sequence 39, Appl	
30	128.5	5.9	4522	13	US-10-329-624-39	Sequence 39, Appl	
31	128	5.9	1830	17	US-10-437-963-61685	Sequence 61685, A	
32	128	5.9	2619	15	US-10-032-585-6421	Sequence 6421, Ap	
33	127.5	5.9	3336	9	US-09-815-242-8857	Sequence 8857, Ap	
34	127	5.9	1695	15	US-10-032-585-6025	Sequence 6025, Ap	
C	35	127	5.9	3628	17	US-10-437-963-22862	Sequence 22862, A
36	126.5	5.9	1191	13	US-10-425-114-4557	Sequence 4557, Ap	
37	126.5	5.9	1464	9	US-09-938-842A-2682	Sequence 2682, Ap	
38	126.5	5.9	1464	11	US-09-938-842A-2682	Sequence 2682, Ap	
39	126.5	5.9	1998	13	US-10-425-114-16858	Sequence 16858, A	
40	126.5	5.9	3825	13	US-10-282-122A-7653	Sequence 7653, Ap	
41	126.5	5.9	4616	17	US-10-437-963-95869	Sequence 95869, A	
42	126	5.8	2199	15	US-10-032-585-6546	Sequence 6546, Ap	
43	125.5	5.8	798	17	US-10-467-243-13	Sequence 13, Appl	
44	125.5	5.8	861	15	US-10-097-340-230	Sequence 230, App	
45	125.5	5.8	861	15	US-10-171-311-175	Sequence 175, App	

```
Alignment Scores:
Pred. No.: 1,08e-233 Length: 2148
Score: 2162.00 Matches: 409
Percent Similarity: 100.00% Conservativeness: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-09-813-329-6 (1-409) x US-09-813-329-5 (1-2148)
QY 1 MetThrAlaGluThrLeuLysProPheIleThrProThrSerAlaAsnAspGlyPhe 20
Db 634 ATGACTGCGGAGACCTCAAGCGTTTATAACGCCAACGAGTGCCAAAGATGATGTTT 693
QY 21 ProAlaLysAlaThrSerThrAlaThrAlaGlnArgArgThrArgGlnLeuIleProLeu 40
Db 694 CCGGCCAAAGCGACGACGACGCGCGACCGCCGAGCGACCGCCGCGACGCTGATCCCGCTG 753
QY 41 ValLeuGlyPheIleGlyLeuGluValValAlaIleLeuAlaLeuThrIleTrpGln 60
Db 754 GTTTTGGGGTTTCATCGGTCTGGGGCTGGTGTGTCATTCGCACTAACCGATCTGGCAG 813
QY 61 ThrThrArgValSerHisLeuAspLysGluLeuLysSerLeuLysArgValValAspAsn 80
Db 814 ACAACGCGTGTATCGCATCTGCACAAAGGAGCTGAAGAGCCTGAAGCGAGTCTGTCGATAAT 873
QY 81 LeuGlnGlnArgLeuGlyLeuAsnTyrLeuAspGluPheAspGluPheGlnLysGluTyr 100
Db 874 CTCACGACGCGTTTGGGCATAACTATCTGACGAGTTCGACGAGTTCGCAAAAGGAGTAC 933
QY 101 GluAsnAlaLeuIleAspTyrProLysLysValAspGlyLeuThrAspGluGluAspAsp 120
Db 934 GAGAAATGCGCTCATCGACTATCCAAAAGAGTGGATGGCTCACGGATGAGGAGGACGAC 993
QY 121 AspAspGlyAspGlyLeuAspSerIleAlaAspAspGluAspAspValSerTyrSer 140
Db 994 GACGATGCGGATGCTGATTCGATTCGATTCGACGACGACGACGACGACGATGATGATGAC 1053
QY 141 SerValAspAspValGlyAlaAspTyrGluAspTyrThrAspMetLeuAsnLysLeuAsn 160
Db 1054 TCTGTGGATGATGTTGGCGAGACTACGAGGACTACACCCGATATGTTAAATAAACAAC 1113
QY 161 AsnAlaIleThrGlyThrThrProThrSerGluThrThrAlaGluGlyGluGlyGluThr 180
Db 1114 AATGCACATACCGCACCCACCCACATCTGAGACCACTGCTGAGGGCGAGGCGAGACG 1173
QY 181 AspSerAlaSerSerAlaSerAsnAspAspAsnValPheAspAspPheThrSerTyrAsn 200
Db 1174 GACAGTGCATCTCAGCGCTCAAAATGATGACAAATGTTTGGATGACATTTACGAGTACAAT 1233
QY 201 AlaHisLysLysLysGlnGluArgLysSerArgSerIleAlaAspValArgAsnGluGlu 220
Db 1234 GCCCACAARAAGACGAGGAGAGAAATCTCGCTCGATTCGCCATGTACCAATGAGGAG 1293
QY 221 GlnAsnIleGlnGlyAsnHisThrGluLeuGlnGlyLysSerSerAsnGluAlaThrSer 240
Db 1294 CAGAAATATCAAGGAAATCACACAGAGCTTCAGGAAAGTCATCCAATGAGGCAACTTCC 1353
QY 241 LysGluSerProAlaProLeuHisHisArgArgArgMetHisSerArgHisArgHisLeu 260
Db 1354 AAAGAGAGCCCTGCACCACTTCACCCGCTCGAGAAATGCAATTCGCGCCATCGCCACCTC 1413
QY 261 LeuValArgLysAlaArgSerGluAspSerArgProAlaAlaHisPheHisLeuSerSer 280
Db 1414 CTAGTCGCAAGACGAGATCCGAGGACTCGAGGCCAGAGCCCATTTCCATTTGACGACG 1473
QY 281 ArgArgArgHisGlnGlySerMetGlyTyrHisGlyAspMetTyrIleGlyAsnAspAsn 300
Db 1474 AGCGCGGTCCCAAGAGAGTATGGGCTACCATGGAGATATGTACATAGAAATGATAAC 1533
QY 301 GluArgAsnSerTyrGlnGlyHisPheGlnThrArgAspGlyValLeuThrValThrAsn 320
Db 301 |||||

US-09-813-329-6 (1-409) x US-09-813-329-3 (1-1221)
QY 1 MetThrAlaGluThrLeuLysProPheIleThrProThrSerAlaAsnAspGlyPhe 20
Db 1 ATGACTGCGGAGACCTCAAGCGTTTATAACGCCAACGAGTGCACGATGATGTTT 60
QY 21 ProAlaLysAlaThrSerThrAlaThrAlaGlnArgArgThrArgGlnLeuIleProLeu 40
Db 61 CCGGCCAAAGCGACGACGACGCGCGACCGCCGAGCGACCGCCGAGCTGATCCCGCTG 120
QY 41 ValLeuGlyPheIleGlyLeuGluValValAlaIleLeuAlaLeuThrIleTrpGln 60
Db 121 GTTTTGGGGTTTCATCGGTCTGGGGCTGGTGTGTCATTCGCACTAACCGATCTGGCAG 180
QY 61 ThrThrArgValSerHisLeuAspLysGluLeuLysSerLeuLysArgValValAspAsn 80
Db 181 ACAACGCGTGTATCGCATCTGCACAAAGGAGCTGAAGAGCCTGAAGCGAGTCTGTCGATAAT 240
QY 81 LeuGlnGlnArgLeuGlyIleAsnTyrLeuAspGluPheAspGluPheGlnLysGluTyr 100
Db 81 |||||

US-09-813-329-3
; Sequence 3, Application US/09813329
; Patent No. US20020012968A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: No. US20020012968A1el Drosophila Tumor Necrosis Factor Class Molec
; FILE REFERENCE: D0016.np
; CURRENT APPLICATION NUMBER: US/09/813,329
; PRIOR FILING DATE: 2001-03-20
; PRIOR FILING DATE: 2000-03-21
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: Patent version 3.0
; SEQ ID NO 3
; LENGTH: 1221
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1) .. (1218)
US-09-813-329-3
Alignment Scores:
Pred. No.: 4,44e-225 Length: 1221
Score: 2082.50 Matches: 400
Percent Similarity: 96.39% Conservativeness: 0
Best Local Similarity: 96.39% Mismatches: 0
Query Match: 96.39% Indels: 15
DB: 9 Gaps: 2
```



```

Db      241  |||||CTCCAGCAGCGCTTTGGGCATAAACTATCTCGACGAGTTCGACGAGTTCCTCAAAAGGAGTAC 300
Qy      101  GluAsnAlaLeuIleAaspTyrProLysValAaspGlyLeuThrAaspGluGluAasp 120
Db      301  GAGATGCGCCCTCATCGACTATCCAAAAGGTGATGGCCCTCAGGATGAGGAGGACGAC 360
Qy      121  AspAaspGlyAaspGlyLeuAaspSerIleAlaAaspAaspGluAaspAaspValSerTyrSer 140
Db      361  GACGATGCGGATGGTCTGATTCCATTCCGACGACGAGGAGGACGACGAGTTCGATATAGC 420
Qy      141  SerValAaspAaspValGlyAlaAaspTyrGluAaspTyrThrAaspMetLeuAasnLysLeuAasn 160
Db      421  TCTGTGGATGATGTTGGCGCAGACTACGAGGACTACACCGATATGTTAAATAAACTCAAC 480
Qy      161  AsnAlaHisThrGlyThrProThrSerGluThrThrAlaGluGlyGluGlyGluThr 180
Db      481  AATGCACATACCGCACCCACCCACATCTCAGACCACTCTGAGGGCGAGGGCGAGACG 540
Qy      181  AspSerAlaSerSerAlaSerAsnAaspAaspValPheAaspAaspPheThrSerTyrAsn 200
Db      541  GACAGTGATCTCTCAGCGCTCAATGATGACAAATGTGTTTCGATGACTTTTACCGACTACAAT 600
Qy      201  AlaHisLysLysLysGlnGluArgLysSerArgSerIleAlaAaspValArgAasnGluGlu 220
Db      601  GCGCACAAAAAGACGACGAGAGAAAATCTCGCTCGATTGCGGATGTACGCAATGAGGAG 660
Qy      221  GlnAsnIleGlnGlyAasnHisThrGluLeuGlnGluLysSerSerAasnGluAlaThrSer 240
Db      661  CAGAAATATTCAAGGAAATACACAGAGCTTCAGGAAAAGTTCATCAATGAGGCAACTTCC 720
Qy      241  LysGluSerProAlaProLeuHisArgArgMetHisSerArgHisArgHisLeu 260
Db      721  AAAGAG-----AGATGCATTCCCGCCATCGCCACCTC 753
Qy      261  LeuValArgLys-----AlaArgSerGluAaspSerArgProAlaAla 274
Db      754  CTAGTCCCAAGGTGAATCTCTCTTTCAGCCAGATCCGAGGACTCGAGGCCAGCAGCC 813
Qy      275  HisPheHisLeuSerSerArgArgHisGlnGlySerMetGlyTyrHisGlyAaspMet 294
Db      814  CATTTCCACTTGAGCAGCGCGCGCTCACCAGGAAGTATGGGCTACCATGGAGATATG 873
Qy      295  TyrIleGlyAasnAaspAasnGluArgAasnSerTyrGlnGlyHisPheGlnThrArgAaspGly 314
Db      874  TACATAGGAATGATACACGAGAACTCTTATCAGGACACTTTCAAACCGCGATGGC 933
Qy      315  ValLeuThrValThrAasnThrGlyLeuTyrTyrValTyrAlaGlnIleCysTyrAasnAasn 334
Db      934  GTCTTTGACGGTGACCAATACAGGCTATATTACGTATACGCCACGATATGCTACCAACAC 993
Qy      335  SerHisAaspGlnAasnGlyPheIleValPheGlnGlyAaspThrProPheLeuGlnCysLeu 354
Db      994  TCGCACGACCAAGAACGGATTTATCGCTCTTTCAAAGGAGACATCTCCATTCCTCGATGCTTG 1053
Qy      355  AasnThrValProThrAasnMetProHisLysValHisThrCysHisThrSerGlyLeuIle 374
Db      1054  AACACGGTGCCCAACACATCCACATTAAGTGCACACCTGCCACACGAGTGGTCTGATC 1113
Qy      375  HisLeuGluArgAasnGluArgIleHisLeuLysAaspIleHisAasnAaspArgAasnAlaVal 394
Db      1114  CACCTGGAACGAAACGAGAGGATCCATCTGAAGGACATTCACACGATTCGCAATGCGAGTT 1173
Qy      395  LeuArgGluGlyAasnAasnArgSerTyrPheGlyIlePheLysVal 409
Db      1174  CTGCGGGAGGGAAACAAACCGAAGTACTTTGGCATCTTCAAGGTC 1218
```

RESULT 3

```

US-09-813-329-1
; Sequence 1, Application US/09813329
; Patent No. US20020012968A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
```

```

; TITLE OF INVENTION: No. US20020012968A1el Drosophila Tumor Necrosis Factor Class Mole
; FILE OF INVENTION: Variants Thereof
; FILE REFERENCE: D0016.np
; CURRENT APPLICATION NUMBER: US/09/813,329
; CURRENT FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: 60/190,816
; PRIOR FILING DATE: 2000-03-21
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 2166
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (652)..(1878)
US-09-813-329-1
Alignment Scores:
Pred. No.: 1,84e-220 Length: 2166
Score: 2045.00 Matches: 394
Percent Similarity: 96.82% Conservative: 2
Best Local Similarity: 96.33% Mismatches: 13
Query Match: 94.59% Indels: 0
DB: 9 Gaps: 0
US-09-813-329-6 (1-409) x US-09-813-329-1 (1-2166)
Qy      1  MetThrIaGluThrLeuLysProPheIleThrProThrSerAlaAasnAaspGlyPhe 20
Db      652  ATGACTGCCGAGACCCCTCAAGCGCTTTATTAACCCAAACGAGTGCCAAACGATGATGGT 711
Qy      21  ProAlaLysAlaThrSerThrAlaThrAlaGlnArgArgThrArgGlnLeuIleProLeu 40
Db      712  CCGGCCAANACGACACGACGCGACCGCCAGCGACGCCGCCAGCTGATCCCCCTG 771
Qy      41  ValLeuGlyPheIleGlyLeuGlyLeuValValAlaIleLeuAlaLeuThrIleTrpGln 60
Db      772  GTTTTGGGGTTTCATCGCTCTGGGGCTGGTCTGGCCATTCTCGCATTACGATCTGCGAG 831
Qy      61  ThrThrArgValSerHisLeuAaspLysGluLeuLysSerLeuLysArgValValAaspAasn 80
Db      832  ACAACGCGTGTATCGCATCTCGGACAAAGGAGCTGAAGAGCCTGAAGCGAGTCTCGATAAT 891
Qy      81  LeuGlnGlnArgLeuGlyIleAasnTyrLeuAaspGluPheAaspGluPheGlnLysGluTyr 100
Db      892  CTCAGCAGCGCTTTGGGCATAAACTATCTCGACGAGTTCGACGAGTTCCTCAAAAGGAGTAC 951
Qy      101  GluAasnAlaLeuIleAaspTyrProLysValAaspGlyLeuThrAaspGluAaspAasp 120
Db      952  GAGAAATGCCCTCATCGCATATCCAAAAAGGTGGATGGCTCACGGATGAGGAGGACGAC 1011
Qy      121  AspAaspGlyAaspGlyLeuAaspSerIleAlaAaspAaspGluAaspAaspValSerTyrSer 140
Db      1012  GACGATGGCGATGGTCTGGATTTCATTTCGCGACGACGAGGACGACGACGCTTAGCTATAGC 1071
Qy      141  SerValAaspAaspValGlyAlaAaspTyrGluAaspTyrThrAaspMetLeuAasnLysLeuAasn 160
Db      1072  TCTGTGATGATGTTGGCGCAGACTACGAGGACTACACCGATATGTTAAATAAACTCAAC 1131
Qy      161  AsnAlaHisThrGlyThrProThrSerGluThrThrAlaGluGlyGluGlyGluThr 180
Db      1132  AATGCACATACCGCACCCACCCACATCTGAGACCACTCTGAGGGCGAGGGCGAGACG 1191
Qy      181  AspSerAlaSerSerAlaSerAasnAaspAasnValPheAaspAaspPheThrSerTyrAasn 200
Db      1192  GACAGTGCATCTCAGCCTCAAAATGATGACAAATGTGTGTCGATGACTTTTACCGACTCAGAT 1251
Qy      201  AlaHisLysLysLysGlnGluArgLysSerArgSerIleAlaAaspValArgAasnGluGlu 220
Db      1252  GCCCTCAAAAAGACGAGGAGAAAATCTCGCTCGGATTCGCCGATGTACCAATGAGGAG 1311
Qy      221  GlnAasnIleGlnGlyAasnHisThrGluLeuGlnGluLysSerSerAasnGluAlaThrSer 240
```



```
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_14156C.1
US-10-424-599-46034

Alignment Scores:
Pred. No.: 7,88e-06 Length: 1269
Score: 143.00 Matches: 72
Percent Similarity: 35.69% Conservative: 44
Best Local Similarity: 22.15% Mismatches: 119
Query Match: 6.61% Indels: 91
DB: 13 Gaps: 9

US-09-813-329-6 (1-409) x US-10-424-599-46034 (1-1269)
Qy 16 AenAspArgGlyPheProAlaLysAlaThrSerThrAlaThrAlaGlnArgThrArg 35
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
148 AACGCAATGCTTTTACATCGACCGCTCCAAATCAAGTCCACGTATCATCGTCGGTGAA 207
Qy 36 GlnLeuIlePro-----
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
208 TCGATTTTGAACGGCCAGCATCAGCGCTGCAGAAAACCTGGGGGAAGCGAAGGAGG 267
Qy 40 -----LeuValLeuGlyPheIleGlyLeuGlyLeuValValAlaIleLeuAla 55
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
268 CTGCTAACTTGGCTGCTTTTCTCCATCTCTCGGTGCTCGTGTTCGATCGAATTTTTTT 327
Qy 56 LeuThrIleTprGlnThrThrArgValSer-----
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
328 CTCGGTTGTTGGGAACACAGAGTTTCGAAATCCCGTGGAGAACGTAAATCTTTTGG 387
Qy 66 HisLeuAspLysGluLeuLysSerLeuLysArgValValAspAsnLeuGlnGlnArgLeu 85
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
388 TTTCTGGATTTGAAGATGAAGCCACAAAGAAATGAGCCACGATGTGGAGAAATGAG 447
Qy 86 GlyIleAsnTyrLeuAspGluPheAspGluPheGlnLysGluTyrGluAsnAlaLeuIle 105
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
448 GTCACGGTCTTGGTTGATGAGGAGGAGGAGGAGAAATACGAGTCCGAGGTGAAGTAGTA 507
Qy 106 AspTyrProLysValAspGlyLeuThrAspGluGluAspAspAspGlyAspGly 125
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
508 GAC-----GACGGAGAGCAGATGAAGACGACGACGACGACGATGACGAT 552
Qy 126 LeuAspSerIleAlaAspAspGluAspAspValSerTyrSerSerValAspAspVal 145
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
553 GACGACGACATGAAGACGAGGAGGAGCAGCAT-----GAAGATGATGCT 597
Qy 146 -----GlyAlaAspTyrGluAspTyrThrAspMetLeuAsnLysLeuAsnAlaHis 163
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
598 CCCGACGCGGTGACGACGACGATGAAGACGACGACGACGACGATGACGAT 630
Qy 164 ThrGlyThrThrProThrSerGluThrAlaGluGlyGluGlyGluThrAspSerAla 183
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
631 -----GAAAGAAAGAAAGAGTGTCCACGCGCGT 660
Qy 184 SerSerAlaSerAsnAspAsnValPheAspAspPheThrSerTyrAsnAlaHisLys 203
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
661 GGAGAGCTCGACGACGATACGATCAGCAGCAGC----- 696
Qy 204 LysLysGlnGluArgLysSerArgSerIleAlaAspValArgAsnGluGluGlnAsnIle 223
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
697 -----GAAAGATGAAGACGAGGAGGATGAGGAGGAG 726
Qy 224 GlnGlyAsnHisThrGluLeu-GlnGluLysSerSerAsnGluAlaThrSerLysGluSe 243
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
727 CAAGGGGAAGAGCGGATTTGGGAACAGACAGTATCTTCCGCCCTCTAGTCACGTG 781
Qy 243 rProAlaProLeuHisHisArgArgArgMetHis-----SerArgHisArgHisLe 260
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
782 -----CTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAT 825
Qy 260 uLeuValArgLysAlaArgSerGluAspSerProAlaAlaHisPheHisLeuSerSe 280
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```



## RESULT 9

```
RESULT 10
US-09-729-658B-14
; Sequence 14, Application US/09729658B
; Publication No. US20030023991A1
; GENERAL INFORMATION:
; APPLICANT: Zonana et al.
; TITLE OF INVENTION: Hypohydrotic ectodermal dysplasia genes and proteins
; FILE REFERENCE: 55924
; CURRENT APPLICATION NUMBER: US/09/729,658B
; CURRENT FILING DATE: 2000-12-04
; PRIOR APPLICATION NUMBER: 09/342,681
; PRIOR FILING DATE: 1999-06-29
; PRIOR APPLICATION NUMBER: 60/092,279
; PRIOR FILING DATE: 1998-07-09
; PRIOR APPLICATION NUMBER: 60/112,366
; PRIOR FILING DATE: 1998-12-15
; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
; LENGTH: 1176
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1176)
US-09-729-658B-14

Alignment Scores:
Pred. No.: 2,58e-05 Length: 1176
Score: 138.00 Matches: 96
Percent Similarity: 35.92% Conservative: 57
Best Local Similarity: 22.54% Mismatches: 147
Query Match: 6.38% Indels: 127
DB: 10 Gaps: 18

US-09-813-329-6 (1-409) x US-09-729-658B-14 (1-1176)
QY 35 ArgGlnLeuIlePro-LeuValLeuGlyPheIleGlyLeuValValAlaIleLe 54
DB 108 AGGAACAGCTGCTCTTCCTGGGTTCTTTGGGCTCTCGTGGCCCTCCACCTGCT 167
QY 54 uAlaLeuThrIleTrpGlnThrThrArgValSerHisLeuAspLysGluLeuLysSerLe 74
DB 168 GACGTTGGTGCTGTAC-----CTAGAGTTGGCTCGGA 200
QY 74 uLysArgValValAspAsnLeuGlnArgLeuGlyLysLeuAsnTyrLeuAspGluPheAs 94
DB 201 GTTCGGCGGGAACGTCGGAGCCGAGTCCCGCTTGGC----- 237
QY 94 pGluPheGlnLysGluTyrGluAsnAlaLeuIleAspTyrProLysLysValAspGlyLe 114
DB 237 ----- 237
QY 114 uThrAspGluGluAspAspAspGlyAspGly-----LeuAs 127
DB 238 -----GGCTCGGGACCCCTGGCACCTCTGGCACCTTAAG 272
QY 127 pSerIleAlaAsp---AspGluAspAspValSerTyrSerSerValAspAspValG1 146
DB 273 CAGCCTCGGTGGCTCGACCTGACAGCCCATCACC----- 309
QY 146 yAlaAspTyrGluAspTyrThrAspMetLeuAsnLysLeuAsnAlaHisThrGly-- 165
DB 310 -----AGTCACCTTGGCA 323
QY 166 -ThrThrProThrSerGluThrThrAlaGluGlyGluGlyGluThrAspSerAlaSerSe 185
DB 324 GCCGTCACTTAAGCAGCAGCATTTGGAAACCGGAGAACCGGACTCTGACTCCCA 383
QY 185 rAlaSerAsnAsp-----AspAsnValPheAspAspPheThrSerTyrAsnAl 201
DB 384 GGACGGCACACGATGGCCCTATTGAATTTCTTCTCCGTGATGAAGAAAGCCACTCTGA 443
```

```
QY 201 aHisLysLysLysGlnGluArg-----LysSerArgSerIleAlaAsp-- 215
DB 444 AGAAGAAAGTAGCGGTTCGCCGCAATAAAAGAACCAATGAAGAGCAGATGG 503
QY 216 ---ValArgAsnGluGluGlnAsnIleGlnGlyAsnHisThrGluLeuGlnGluLysSe 234
DB 504 CCAGTTTAAACAA-GAAAAGGGAAAGAAAGCAGGACCTCTCGACCAATGGCCCTC 562
QY 234 rSerAsnGluAlaThrSerLysGluSerProLysProLysHisHisArgArgMethHi 254
DB 563 CAGGACCCCGCAGGACCTCCAGGACCCCGGACCCCGGAAATTCAGGGATTCCTGGAA 622
QY 254 sSerArgHisArg---HisLeuLeuValArgLysAlaArgSerGluAspSer-- 270
DB 623 TTCAGGAACAACCTGTATTATGGGACCACTGGTCTCTCCAGGTCTCTCTGCTCAAGGAC 682
QY 271 -----ArgProAl 273
DB 683 CCCCTGGCTCCAGGACCTTCTGGTGTGTGATATAAGCTGGAACCTCGAAGAACCAAC 742
QY 273 a-AlaHisPheHisLeuSerSerArgArgHisGlnGlySer---MetGlyTyrHisG 292
DB 743 CAGCTGTGGTGCATCTACAGGGC-----CAAGGGTCAGCAATTCAAGTCAACA 790
QY 292 LyAspMetTyrIleGly-----AsnAspAsnGluArg-----AsnSerTyrGlnG 307
DB 791 ATGATCTTTCAGGTGGAGTGCTCAATGACTGTCTGTCATCATATGAACCCCAAGGTGT 850
QY 307 LyHisPheGlnThrArgAspGlyValLeuThrValThrAsnThrGlyLeuTyrTyrValT 327
DB 851 TTAAGCTACATCCCGCAGCGGGGAGCTGGAGGTACTGGTGACCGCACCTACTTCACT 910
QY 327 YrAlaGln-----IleCysTyrAsnAsnSerHisAspGlnAsnGlyPheIleValPheG 345
DB 911 ATAGTCAGGTAGAAGTATACATCACTTTCCTGACTTTCAGTGTGAGTGGTGGTGG 970
QY 345 LngLysPthrProPheLeuGlnCysLeuAsnThrValProThrAsnMetProHisLysEv 365
DB 971 TGGATGAGAGCCCTTCTTCAGTGCACACCGCAGCATCGAGACGGGCAAGCAAC--T 1027
QY 365 aHisThrCysHisThrSerGlyLeuIleHisLeuGluArgAsnGluArgIleHisLeuL 385
DB 1028 ACAACACTGTCTATACCGCAGCGCTGCTCTCTCAAGCCCGCGCAGAAAGATCGCGTCA 1087
QY 385 yAspIleHisAsnAspArgAsnAlaValLeuArgGluGlyAsnAsnArgSerTyrPheG 405
DB 1088 AGATGTGCACGCTGACATCTCC-----ATCAACATGAGCAAGCACACACGTTCTTTG 1141
QY 405 LyIlePheLysVal 409
DB 1142 GGGCCATCAGGCTG 1155

RESULT 11
US-09-729-658B-1
; Sequence 1, Application US/09729658B
; Publication No. US20030023991A1
; GENERAL INFORMATION:
; APPLICANT: Zonana et al.
; TITLE OF INVENTION: Hypohydrotic ectodermal dysplasia genes and proteins
; FILE REFERENCE: 55924
; CURRENT APPLICATION NUMBER: US/09/729,658B
; CURRENT FILING DATE: 2000-12-04
; PRIOR APPLICATION NUMBER: 09/342,681
; PRIOR FILING DATE: 1999-06-29
; PRIOR APPLICATION NUMBER: 60/092,279
; PRIOR FILING DATE: 1998-07-09
; PRIOR APPLICATION NUMBER: 60/112,366
; PRIOR FILING DATE: 1998-12-15
; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1574
; TYPE: DNA
```





```

; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (4545)..(4545)
; OTHER INFORMATION: n equals a, g, c or t
US-10-202-062-41

Alignment Scores:
Pred. No.: 0.000255 Length: 5307
Score: 138.00 Matches: 96
Percent Similarity: 35.92% Conservative: 57
Best Local Similarity: 22.54% Mismatches: 147
Query Match: 6.38% Indels: 127
DB: 13 Gaps: 18

US-09-813-329-6 (1-409) x US-10-202-062-41 (1-5307)

Qy 35 ArgGlnLeuPro-LeuValLeuGlyPheIleGlyLeuGlyValValAlaIleLe 54
Db 350 AGGAACAGCTGCTGCTCTCTCTGGGTCTTTTGGCTCTCGCTGGCCCTCCACCTGCT 409
Qy 54 uAlaLeuThrIleTrpGlnThrThrArgValSerHisLeuAspLysGluLeuLysSerLe 74
Db 410 GACCTTGCTGCTGCTAC-----CTAGAGTTGGCTCGGA 442
Qy 74 uLysArgValValAspAsnLeuGlnArgLeuGlyLeuAsnTyrIleAsnTyrLeuAspGluPheAs 94
Db 443 GTTCGGGGGAGACGTGGAGCGGAGTCCGCGCTTGGC----- 479
Qy 94 pGluPheGlnLysGluTyrGluAsnAlaLeuIleAspTyrProLysLysValAspGlyLe 114
Db 479 ----- 479
Qy 114 uThrAspGluGluAspAspAspGlyAspGly-----LeuAs 127
Db 480 -----GGCTCGGGGACCCCTGGCACCCTCGGCACCCCTAAG 514
Qy 127 pSerIleAlaAsp---AspGluAspAspAspValSerTyrSerSerValAspAspValG1 146
Db 515 CAGCTCGTGGGCTGACCTGACAGCCCATCACC----- 551
Qy 146 yAlaAspTyrGluAspTyrThrAspMetLeuAsnLysLeuAsnAsnAlaHisThrGly-- 165
Db 552 -----AGTCACCTTGGCA 565
Qy 166 -ThrThrProThrSerGluThrThrAlaGluGlyGluGlyGluThrAspSerAlaSerSe 185
Db 566 GCCGTCACTAAGCAGCAGCATTTGGAAACCGGGAGAACCGCACTCCCACTCTGACTCCCA 625
Qy 185 rAlaSerAsnAsp-----AspAsnValPheAspAspPheThrSerTyrAsnAl 201
Db 626 GGACGGGACACAGATGCCCTATTGAATTTCTTCTCCCTGATGAAGCCATACCTGA 685
Qy 201 aHisLysLysGlnGluArg-----LysSerArgSerIleAlaAsp-- 215
Db 686 AGAAGAAAGTAGGCGTGTTCGCCCAATAAAGAACGAAAGCAATGAAGCAGCAGATGG 745
Qy 216 ----ValArgAsnGluGlnAsnIleGlnGlyAsnHisThrGluLeuGlnGlyLysSe 234
Db 746 CCCAGTTAAAAACAA--GAAAAAGGAAAGAAAGCAGACCTCTCGGACCCCAATGGCCCTC 804
Qy 234 rSerAsnGluAlaThrSerLysGluSerProAlaProLeuHisHisArgArgMetHi 254
Db 805 CAGGACCCCGAGGACCTCCAGGACCCCGAGGATTCAGGATTCCTGGAA 864
Qy 254 sSerArgHisArg---HisLeuLeuValArgLysAlaArgSerGluAspSer----- 270
Db 865 TTCCAGGAACAACATGTTATGAGCACCATGCTGCTCCAGGTCTCTCTGGCTCAAGGAC 924
Qy 271 -----ArgProAl 273
Db 925 CCCCTGGCTCCAGGACCTTCTGGTGTGCTGATAAGCTGGAACCTCGAATAACCCAGC 984
Qy 273 a-AlaHisPheHisLeuSerSerArgArgHisGlnGlySer---MetGlyTyrHisG 292

```

## RESULT 13

```

US-10-272-411-6
; Sequence 6, Application US/10272411
; Publication No. US20030100068A1
; GENERAL INFORMATION:
; APPLICANT: Barnes Jewish Hospital
; APPLICANT: Lam, Jonathan
; APPLICANT: Ross, F. Patrick
; APPLICANT: Teitelbaum, Steven
; TITLE OF INVENTION: RANKL MIMICS AND USES THEREOF
; FILE REFERENCE: 60019620-0202
; CURRENT APPLICATION NUMBER: US/10/272,411
; CURRENT FILING DATE: 2002-10-15
; PRIOR APPLICATION NUMBER: 60/329,393
; PRIOR FILING DATE: 2001-10-15
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 5307
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (4242)..(4242)
; OTHER INFORMATION: n = a, t, c or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (4471)..(4471)
; OTHER INFORMATION: n = a, t, c or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (4523)..(4523)
; OTHER INFORMATION: n = a, t, c or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (4529)..(4529)
; OTHER INFORMATION: n = a, t, c or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (4531)..(4531)
; OTHER INFORMATION: n = a, t, c or g

```



<hr/>					
:	NAME/KEY:	misc feature			
:	LOCATION:	(4545)..(4545)			
:	OTHER INFORMATION:	n = a, t, c or g			
:	PUBLICATION INFORMATION:				
:	DATABASE ACCESSION NUMBER:	NCBI/NM_001399.1			
:	DATABASE ENTRY DATE:	2000-10-31			
:	RELEVANT RESIDUES:	(1)..(5307)			
:	PUBLICATION INFORMATION:				
:	DATABASE ACCESSION NUMBER:	NCBI/AFO40628.1			
:	DATABASE ENTRY DATE:	1998-11-13			
:	RELEVANT RESIDUES:	(1)..(5307)			
:	PUBLICATION INFORMATION:				
:	DATABASE ACCESSION NUMBER:	NCBI/AFO61189.1			
:	DATABASE ENTRY DATE:	1998-11-14			
:	RELEVANT RESIDUES:	(1)..(5307)			
<hr/>					
US-10-272-411-6					
Alignment Scores:					
	Pred. No.:	0.000255	Length:	5307	
	Score:	138.00	Matches:	96	
	Percent Similarity:	35.92%	Conservative:	57	
	Best Local Similarity:	22.54%	Mismatches:	147	
	Query Match:	6.38%	Indels:	127	
	DB:	15	Gaps:	18	
<hr/>					
US-09-813-329-6 (1-409) x US-10-272-411-6 (1-5307)					
Qy	35	ArgGlnLeuIlePro-LeuValLeuGlyPheIleGlyLeuValValAlaIleLe 54			
Db	350	AGGAACACGTGCCTCCTTCCTGGGTTTCTTTGGGCTCTCGCTGGCCCTCACCTGTGT 409			
Qy	54	uAlaLeuThrIleTrpGlnThrThrArgValSerHisLeuAspLysGluLeuLysSerLe 74			
Db	410	GACCTTGTCGTCTATC-----CTAGAGTTGGCTCGGA 442			
Qy	74	uLysArgValValAspAsnLeuGlnArgLeuGlyIleasnTyrlEuAspGluPheAs 94			
Db	443	GTTGCGCGCGGAACGTGGAGCGAGTCGCCGCTTGGC----- 479			
Qy	94	pGluPheGlnLysGluTyrlGluAsnAlaLeuIleAspTyrlProLysLysValAspGlyLe 114			
Db	479	----- 479			
Qy	114	uThrAspGluGluAspAspAspGlyAspGly-----LeuAs 127			
Db	480	-----GGCTGGGCACCCTGGCACCTCTGGCACCCCTAAG 514			
Qy	127	pSerlleAlaAsp---AspGluAspAspValSerTyrlSerSerValAspAspValGl 146			
Db	515	CAGCCTCGTGCCCTCGACCCCTGACAGCCCATCACC----- 551			
Qy	146	yAlaAspTyrlGluAspTyrlAspMetLeuAsnLysLeuAsnAlaHisThrGly-- 165			
Db	552	-----AGTCACTTGGCCA 565			
Qy	166	-ThrThrProThrSerGluThrThrAlaGluGlyGluGlyuThrAspSerAlaSerSe 185			
Db	566	GCGCTCACCTAAAGCAGCATTGGAACCGGAGAGCCGCACTCCACTCTGACTCCCCA 625			
Qy	185	rAlaSerAsnAsp-----AspAsnValPheAspAspPheThrSerTyrlAsnAl 201			
Db	626	GGACGGGCACAGATGGCCCTATTGAATTTCTTCTCCCTGATGAAGAAGCCATACTCTGA 685			
Qy	201	aHisLysLysLysGlnGluArg-----LysSerArgSerillealaAsp-- 215			
Db	686	AGAAAGAAATAGCGTGTTTCGCCCGCAATAAAGAAAGCAAAGCAATGAAGGAGCATGG 745			
Qy	216	----ValArgAsnGluGluGlnAsnIleGlnGlyAsnHisThrGluLeuGlnGluLysSe 234			
Db	746	CCCAGTTAAAACA-A-GAAAAAGGGAAGAAGCAGGACCTCTGGACCAATGGCCCTC 804			
Qy	234	rSerAsnGluAlaThrSerLysGluSerProAlaProLeuHiHisArgHargMethI 254			

```

; LOCATION: (4523)..(4523)
; OTHER INFORMATION: n equals a, c, g or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (4529)..(4529)
; OTHER INFORMATION: n equals a, c, g or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (4531)..(4531)
; OTHER INFORMATION: n equals a, c, g or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (4545)..(4545)
; OTHER INFORMATION: n equals a, c, g or t
US-10-218-547-41

Alignment Scores:
Pred. No.: 0.000255 Length: 5307
Score: 138.00 Matches: 96
Percent Similarity: 35.92% Conservative: 57
Best Local Similarity: 22.54% Mismatches: 147
Query Match: 6.38% Indels: 127
Db: 15 Gaps: 18

US-09-813-329-6 (1-409) x US-10-218-547-41 (1-5307)

QY 35 ArgGlnLeuIlePro-LeuValLeuGlyPheIleGlyLeuValValAlaIleLe 54
Db 350 AGGGAACAGCTGCTTCCTCGGGTTCTTGGCCCTCTCGCTGGCCCTCCACCTGCT 409
QY 54 uAlaLeuThrIleTrpGlnThrThrArgValSerHisLeuAspLysGluLeuLysSerLe 74
Db 410 GACGTTGTGCTGCTAC-----CTAGAGTTGCGCTCGGA 442
QY 74 uLysArgValValAspAsnLeuGlnArgLeuGlyIleAsnTyrLeuAspGluPheAs 94
Db 443 GTTGGCGGGAACGCTGGAGCCGAGTCCGCCCTTGGC----- 479
QY 94 pGluPheGlnLysGluTyrGluAsnAlaLeuIleAspTyrProLysLysValAspGlyLe 114
Db 479 ----- 479
QY 114 uThrAspGluGluAspAspAspGlyAspGly-----LeuAs 127
Db 480 -----GGCTCGGCACCCCTGGCACCCTCTGGCACCCTAAG 514
QY 127 pSerIleAlaAsp---AspGluAspAspAspValSerTyrSerSerValAspAspValGI 146
Db 515 CAGCCTCGGTGGCTCGACCCCTGACAGCCCATCACC----- 551
QY 146 yAlaAspTyrGluAspTyrThrAspMetLeuAsnLysLeuAsnAsnAlaHisThrGly-- 165
Db 552 -----AGTCACCTTGGGCA 565
QY 166 -ThrThrProThrSerGluThrThrAlaGluGlyGluGlyGluThrAspSerAlaSerse 185
Db 566 GCCCTCACCTAAGCAGCAGCATTTGGAAACCGGAGAGCCGCACTCCCACTCTGACTCCCA 625
QY 185 rAlaSerAsnAsp-----AspAsnValPheAspAspPheThrSerTyrAsnAl 201
Db 626 GGACGGGCACAGATGCCCTATTGAATTTCTTCTCCCTGATGAAAGCCATCTGTA 685
QY 201 aHisLysLysLysGlnGluArg-----LysSerArgSerIleAlaAsp-- 215
Db 686 AGAAGAAAGTAGGGTGTTCGCCCATATAAAGAGCAAAAGCAATGAAGACAGATGG 745
QY 216 ----ValArgAsnGluGluGlnAsnIleGlnGlyAsnHisThrGluLeuGlnGluLysSe 234
Db 746 CCCAGTTAAAAACAA--GAAAAAGGGAAGAAAGCAGGACCTCTGGACCAATGGCCCTC 804
QY 234 rSerAsnGluAlaThrSerLysGluSerProAlaProLeuHisHisArgArgMetHi 254
Db 805 CAGGACCCCGAGGACCTCCAGGACCCCGAGGACCCCGAGGATTCAGGGATTCTCTGGAA 864

254 sSerArgHisArg--HisLeuLeuValArgLysAlaArgSerGluAspSer----- 270
865 TTCAGGAACAACACTGTATGGGACCACTGGTCTCCAGGTCTCTCTGGTCTCTCAAGGAC 924
271 -----ArgProAl 273
925 CCCTGGCTCCAGGACCTTCTGGTGTCTGCTGATAAAGCTGGAACCTCAGAAACACAGC 984
273 a-AlaHisPheHisLeuSerSerArgArgHisGlnGlySer---MetGlyTyrHisG 292
985 CAGCTGTGGTGCATCTACAGGC-----CAAGGGTCAGCAATTCAAGTCAAGA 1032
292 LyAspMetTyrIleGly-----AsnAspAsnGluArg-----AsnSerTyrGlnG 307
1033 ATGATCTTTCAGGTGGAGTGCTCAATGACTGTCTCGCATCACTATGAACCCCAAGGTGT 1092
307 LyHisPheGlnThrArgAspGlyValLeuThrValThrAsnThrGlyLeuTyrTyrValT 327
1093 TTAAGCTACATCCCGCAGCGGGGAGCTGGAGGTACTGGTGGACGCACTACTTCTACTCT 1152
327 YrAlaGln-----IleCysTyrAsnAsnSerHisAspGlnAsnGlyPheIleValPheG 345
1153 ATAGTCAGGTAGAAGTATATACATCACTTCACTGACTTTCGCACTATGAGGTGGTGG 1212
345 LngLysPheThrProPheLeuGlnCysLeuAsnThrValProThrAsnMetProHisLysV 365
1213 TGGATGAAGAGCCCTTCCTGCGAGTGCAACGACGATCGAGCGGCAAGCAACCAAC---T 1269
365 alHisThrCysHisThrSerGlyLeuIleHisLeuGluArgAsnGluArgIleHisLeuL 385
1270 ACAACACTTGTATACCGCAGCGGCTGCTCTCTCAAGSCCGCGCAGAGATCGCGTCA 1329
385 yAspIleHisAsnAspArgAsnAlaValLeuArgGluGlyAsnAsnAsnArgSerTyrPheG 405
1330 AGATGTGTCAGCTGACATCTCC-----ATCAACATGACGAAGCACACCACTGTTCTTTG 1383
405 LyIlePheLysVal 409
1384 GGCCCATCAGGCTG 1397

RESULT 15
US-10-272-328A-6
; Sequence 6, Application US/10272328A
; Publication No. US20030109444A1
; GENERAL INFORMATION:
; APPLICANT: Barnes Jewish Hospital
; APPLICANT: Lam, Jonathan
; APPLICANT: Ross, F. Patrick
; APPLICANT: Teitelbaum, Steven
; TITLE OF INVENTION: RANKL MIMICS AND USES THEREOF
; FILE REFERENCE: 60019620-0206
; CURRENT APPLICATION NUMBER: US/10/272,328A
; PRIOR FILING DATE: 2003-01-24
; PRIOR APPLICATION NUMBER: 60/329,393
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 5307
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (4242)..(4242)
; OTHER INFORMATION: n = a, t, c or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (4471)..(4471)
; OTHER INFORMATION: n = a, t, c or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (4523)..(4523)

```





GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: September 15, 2004, 05:12:07 ; Search time 4190 Seconds  
(without alignments)  
2914.948 Million cell updates/sec

Title: US-09-813-329-6  
Perfect score: 2162  
Sequence: 1 MTATLTKPFTTPTSDNGF.....DRNAVLREGNRSYFGIFKV 409

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 27513289 seqs, 14931090276 residues  
Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n.model -DEV=xlp  
-Q=/cgn2\_1/USPTO\_spool\_P/US09813329/runat\_13092004\_102753\_21680/app\_query.fasta\_1.583  
-DB=EST -QFMT=fascap -SUFFIX=rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US09813329 @CGN 1 1 6425 @runat\_13092004\_102753\_21680 -NCPU=6 -ICPU=3  
-NO MMAP -LARGQUERY -NEG\_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1:	em_estba.*	EST.*
2:	em_esthum.*	
3:	em_estin.*	
4:	em_estmu.*	
5:	em_estov.*	
6:	em_estpl.*	
7:	em_estro.*	
8:	em_hcc.*	
9:	gb_est1.*	
10:	gb_est2.*	
11:	gb_hcc.*	
12:	gb_est3.*	
13:	gb_est4.*	
14:	gb_est5.*	
15:	em_estfun.*	
16:	em_estom.*	
17:	em_gss_hum.*	
18:	em_gss_inv.*	
19:	em_gss_pln.*	
20:	em_gss_vrt.*	
21:	em_gss_fun.*	
22:	em_gss_mam.*	
23:	em_gss_mus.*	
24:	em_gss_pro.*	
25:	em_gss_rtd.*	
26:	em_gss_phg.*	
27:	em_gss_vrl.*	
28:	gb_gss1.*	

29: gb\_gss2.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	843	39.0	686	12	BG636194	BG636194 SD13795.5
2	829	38.3	495	12	B1636564	B1636564 SD18286.5
3	480	22.2	508	14	CK135985	CK135985 SD18286.3
C 4	466	21.6	532	14	CK135525	CK135525 RH51659.3
C 5	397	18.4	612	13	BX560278	BX560278 BX560278
6	206.5	9.6	539	13	BX553466	BX553466 BX553466
7	193.5	9.0	526	9	A1253821	A1253821 AIMS-P.mo
8	172	8.0	1101	29	CNS017MQ	AL108596 Drosophil
9	170	7.9	917	29	CNS07DXT	AL441063 T3 end of
10	151	7.0	709	13	BU373036	BU373036 603812126
11	150	6.9	965	14	CF590046	CF590046 AGENCOURT
12	145	6.7	746	14	CD851068	CD851068 DHOAU52H0
13	143	6.6	550	12	B1514804	B1514804 BB160016A
14	143	6.6	724	13	BU798281	BU798281 SJF2ACB08
C 15	143	6.6	732	13	BU803616	BU803616 SJFAWE05
16	143	6.6	951	11	AY223179	AY223179 Schistoso
17	143	6.6	1148	14	CD509148	CD509148 CDA94-G04
18	143	6.6	2660	11	BC049115	BC049115 Mus muscu
19	141	6.5	746	13	BC620580	BC620580 BX620580
20	140.5	6.5	1413	11	BC020355	BC020355 Mus muscu
21	140.5	6.5	1414	11	BC029991	BC029991 Mus muscu
22	140.5	6.5	2894	11	BC023777	BC023777 Mus muscu
23	140	6.5	688	29	AY401339	AY401339 Mus muscu
24	140	6.5	703	13	BU445245	BU445245 603764805
25	140	6.5	892	13	BU911456	BU911456 AGENCOURT
26	138.5	6.4	291	13	BX560279	BX560279 BX560279
27	137.5	6.4	670	12	BX269011	BX269011 BX269011
28	137	6.3	906	12	B1250058	B1250058 602995408
29	137	6.3	2812	11	BC033648	BC033648 Homo sapi
30	136	6.3	648	12	BM615338	BM615338 170006871
31	135.5	6.3	800	13	BU219475	BU219475 603756623
32	135	6.2	713	13	BU117025	BU117025 603139378
33	135	6.2	936	11	CNS08SL1	BU026465 Single re
34	134.5	6.2	582	10	BF482404	BF482404 WHE1796 C
35	134.5	6.2	866	12	B1249241	B1249241 602995518
36	134	6.2	814	12	B1250570	B1250570 602995630
37	134	6.2	875	14	CA473339	CA473339 AGENCOURT
38	134	6.2	914	12	B1250061	B1250061 602995411
39	134	6.2	923	13	BU771970	BU771970 SJEELA02
40	133	6.2	615	14	CA855938	CA855938 PFESToac5
41	133	6.2	912	13	BU149401	BU149401 AGENCOURT
C 42	132.5	6.1	580	12	BJ379142	BJ379142 BJ379142
C 43	132.5	6.1	679	12	BM203188	BM203188 C0250E10-
44	132.5	6.1	715	12	BJ617087	BJ617087 BJ617087
45	132.5	6.1	780	12	BI108540	BI108540 602894986

ALIGNMENTS

RESULT 1	BG636194	686 bp	mRNA	linear	EST 23-APR-2001
LOCUS	SD13795.5	prime SD	Drosophila melanogaster	Schneider L2 cell culture	
DEFINITION	pot2 Drosophila melanogaster cDNA clone SD13795 5 similar to CG12919; Fban0012919 located on: 2R 46E1-46E1;: 04/13/2001, mRNA sequence.				
ACCESSION	BG636194				
VERSION	BG636194.1	GI:13763731			
KEYWORDS	EST.				
SOURCE	Drosophila melanogaster (fruit fly)				
ORGANISM	Drosophila melanogaster				
	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;				

Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 Ephydroidea; Drosophilidae; Drosophila.  
 1 (bases 1 to 686)  
**REFERENCE**  
**AUTHORS** Harvey,D., Brokstein,P., Hong,L., Evans-Holm,M., Su,C., Tsang,G.,  
 Lewis,S. and Rubin,G.M.  
**TITLE** BDGP/HMI Drosophila EST Project  
**JOURNAL** Unpublished (2001)  
**COMMENT** Contact: Stapleton, M.  
 BDGP

Lawrence Berkeley National Lab  
 One Cyclotron Rd, Berkeley, CA 94720, USA  
 Fax: 510 486 6798  
 Email: [http://www.fruitfly.org/EST\\_est@fruitfly.berkeley.edu](http://www.fruitfly.org/EST_est@fruitfly.berkeley.edu)  
 hit genomic AB003831: arm:2R [4876906,5137815]  
 estimated-cyto:46B13-46E4; 04/13/2001  
 Plate: SD.137 row: H column: 11  
 High quality sequence stop: 685.

#### FEATURES

Location/Qualifiers  
 1..686  
 /organism="Drosophila melanogaster"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:7227"  
 /clone="SD13795"  
 /lab\_host="DH5-alpha"  
 /clone\_lib="SD Drosophila melanogaster Schneider L2 cell  
 culture pOT2"  
 /note="Vector: pOT2; Site 1: EcoRI; Site 2: XhoI; Sized  
 fractionated cDNAs were directly ligated into pOT2.  
 Plasmid cDNA library."

#### ORIGIN

Alignment Scores:  
 Pred. No.: 5,376-65 Length: 686  
 Score: 843.00 Matches: 153  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 38.99% Indels: 0  
 DB: 12 Gaps: 0  
 US-09-813-329-6 (1-409) x BG636194 (1-686)

QY 257 HisArgHisLeuLeuValArgLysAlaArgSerGluAspSerArgProAlaAlaHisPhe 276  
 Db 2 CATGCCACCTCTCTAGTCCGCAAGCCAGATCCGAGACTCGAGGCCAGAGCCCATTC 61  
 QY 277 HisLeuSerSerArgArgHisGlnGlySerMetGlyTyrHisGlyAspMetTyrIle 296  
 Db 62 CACCTGAGCAGCAGCGCGGCTCACCAAGGAAGTAGGGCTACCATGGAGATATGTACATA 121  
 QY 297 GlyAsnAspAsnGluArgAsnSerTyrGlnGlyHisPheGlnThrArgAspGlyValLeu 316  
 Db 122 GGAATGATAACGAGAGAAACTCTTATCAGGACACATTTCAACGCCGCGATGGCTCTTG 191  
 QY 317 ThrValThrAsnThrGlyLeuTyrTyrValTyrAlaGlnIleCysTyrAsnAsnSerHis 336  
 Db 182 ACGGTGACCAATACAGCCCTATATTAGGTATACGCCAGATATGCTACAACTCGCAC 241  
 QY 337 AspGlnAsnGlyPheIleValPheGlnGlyAspThrProPheLeuGlnCysLeuAsnThr 356  
 Db 242 GACCAGAACGGATTATTCGTCTTCAAGAGAGACACTCCATCTCTGCAGTCTTGAACACG 301  
 QY 357 ValProThrAsnMetProHisLysValHisThrCysHisThrSerGlyLeuIleHisLeu 376  
 Db 302 GTGCCCAACCAATGCCACATAGGTGCACACCTGCCACACGAGTGTCTGTATCCACCTG 361  
 QY 377 GluArgAsnGluArgIleHisLeuLysAspIleHisAsnAspArgAsnAlaValLeuArg 396  
 Db 362 GAACGAACGAGAGATCCATCTGAAGGACATTCACAAACATCGCAATGCAGTCTCGCG 421  
 QY 397 GluGlyAsnAsnArgSerTyrPheGlyIlePheLysVal 409  
 Db 422 GAGGGAACCAACCAAGCTACTTTGGCATCTTCAAGGTG 460

#### RESULT 2

Bi636564  
 LOCUS  
 DEFINITION

Bi636564 495 bp mRNA linear EST 10-SEP-2001  
 SD18286.5prime SD Drosophila melanogaster Schneider L2 cell culture  
 pOT2 Drosophila melanogaster cDNA clone SD18286 5 similar to  
 CG12919: F8an0012919 GO:() located on: 2R 46E1-46E1;: 05/19/2001,  
 mRNA sequence.

#### ACCESSION

Bi636564 GI:15538774  
 VERSION  
 KEYWORDS

#### SOURCE

Drosophila melanogaster (fruit fly)

#### ORGANISM

Drosophila melanogaster  
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 Ephydroidea; Drosophilidae; Drosophila.

#### REFERENCE

1 (bases 1 to 495)  
 Harvey,D., Brokstein,P., Hong,L., Evans-Holm,M., Su,C., Tsang,G.,  
 Lewis,S. and Rubin,G.M.

#### AUTHORS

BDGP/HMI Drosophila EST Project

#### TITLE

Unpublished (2001)

#### JOURNAL

#### COMMENT

Contact: Stapleton, M.  
 BDGP

Lawrence Berkeley National Lab

One Cyclotron Rd, Berkeley, CA 94720, USA

Fax: 510 486 6798

Email: [http://www.fruitfly.org/EST\\_est@fruitfly.berkeley.edu](http://www.fruitfly.org/EST_est@fruitfly.berkeley.edu)

hit genomic AB003831: arm:2R [4876906,5137815]

estimated-cyto:46B13-46E4; 05/19/2001

Plate: SD.182 row: H column: 2

High quality sequence stop: 491.

#### FEATURES

##### source

1..495  
 /organism="Drosophila melanogaster"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:7227"  
 /clone="SD18286"  
 /lab\_host="DH5-alpha"  
 /clone\_lib="SD Drosophila melanogaster Schneider L2 cell  
 culture pOT2"  
 /note="Vector: pOT2; Site 1: EcoRI; Site 2: XhoI; Sized  
 fractionated cDNAs were directly ligated into pOT2.  
 Plasmid cDNA library."

#### ORIGIN

##### Alignment Scores:

Pred. No.: 6,068-64 Length: 495  
 Score: 829.00 Matches: 162  
 Percent Similarity: 99.39% Conservative: 0  
 Best Local Similarity: 99.39% Mismatches: 1  
 Query Match: 38.34% Indels: 1  
 DB: 12 Gaps: 0

US-09-813-329-6 (1-409) x Bi636564 (1-495)

QY 100 TyrGluAsnAlaLeuIleAspTyrProLysLysValAspGlyLeuThrAspGluAsp 119  
 Db 6 TACGAGATGCCCTCATCGACTATCCAA-AAGGTGGATGGCTCACGATGAGGAGAC 64  
 QY 120 AspAspAspGlyAspGlyLeuAspSerIleAlaAspAspGluAspAspValSerTyr 139  
 Db 65 GACGACGATGGCGATGGTCTCGATTCCATTCGGACGACGAGCAGCGCTTAGCTAT 124  
 QY 140 SerSerValAspAspValGlyAlaAspTyrGluAspTyrThrAspMetLeuAsnLysLeu 159  
 Db 125 AGCTCTGTGGATGATGTTGGCGCAGACTACCGGACTACCGATATGTTAAATAAATC 184  
 QY 160 AsnAsnAlaHisThrGlyThrThrProThrSerGluThrThrAlaGluGlyGluGlyGlu 179  
 Db 185 AACATGCACATACACCGCACCCACCGCCCACTCTGAGACCACTGCTGAGGGCGAGCGAG 244  
 QY 180 ThrAspSerAlaSerSerAlaSerAsnAspAspAsnValPheAspAspPheThrSerTyr 199  
 Db 245 ACGGACGTGCATCCTCAGCCTCAAATGATGACAATGTGTTTCGATGACTTTTACCAGCTAC 304

Qy 200 AsnAlaHisLysLysGlnGluArgLysSerArgSerIleAlaAspValArgAsnGlu 219  
 |||||  
 Db 305 AATCCCAAAAAGAACGAGAGAAATCTCGCTCGATTGCCGATGTACGCAATGAG 364  
 |||||  
 Qy 220 GluGlnAenIleGlnGlyAsnHisThrGluLeuGlnGluLysSerSerAsnGluAlaThr 239  
 |||||  
 Db 365 GAGCAGATATTTCAAGGAAATCACACAGAGCTTCAGGAAAGTTCATCCAAATGAGGCAACT 424  
 |||||  
 Qy 240 SerLysGluSerProAlaProLeuHisHisArgArgMetHisSerArgHisArgHis 259  
 |||||  
 Db 425 TCCAAAGAGAGCCCTGCACCACTTACCACCCTCGCAGATGCATTCCTCCGCCATCGCCAC 484  
 |||||  
 Qy 260 LeuLeuVal 262  
 |||||  
 Db 485 CTCCTAGTC 493

RESULT 3  
 CK135985/c  
 LOCUS  
 DEFINITION SD18286.3prime SD Drosophila melanogaster Schneider L2 cell culture  
 POT2 Drosophila melanogaster cDNA clone SD18286 3, mRNA sequence.

CK135985  
 CK135985.1 GI:38627921  
 EST.  
 Drosophila melanogaster (fruit fly)

ORGANISM  
 Drosophila melanogaster  
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 Ephydroidea; Drosophilidae; Drosophila.

REFERENCE  
 1 (bases 1 to 508)  
 Harvey,D., Brokstein,P., Hong,L., Evans-Holm,M., Su,C., Tsang,G.,  
 Lewis,S. and Rubin,G.M.

TITLE  
 JOURNAL  
 COMMENT BDGP/HMI Drosophila EST Project  
 Unpublished (2001)  
 Other ESTs: SD18286.5prime  
 Contact: Stapleton, M.

BDGP  
 Lawrence Berkeley National Lab  
 One Cyclotron Rd, Berkeley, CA 94720, USA  
 Fax: 510 486 6798

Email: [http://www.fruitfly.org/EST\\_est@fruitfly.berkeley.edu](http://www.fruitfly.org/EST_est@fruitfly.berkeley.edu)  
 Based upon the presence of a XhoI site followed by a run of 14 or  
 more T residues at the beginning of the sequence, this clone was  
 polyadenylated. The resulting Poly-T sequence has been removed. hit  
 genomic AE003831: arm:2R [4876906,5137815]  
 estimated-cyto:46B13-46E4: 02/08/2002  
 Plate: SD.182 row: H column: 2  
 High quality sequence stop: 456.

FEATURES  
 source  
 1..508  
 /organism="Drosophila melanogaster"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:7227"  
 /clone="SD18286"  
 /lab\_host="DH5-alpha"  
 /clone\_lib="SD Drosophila melanogaster Schneider L2 cell  
 culture POT2"  
 /note="vector: pOT2; Site 1: EcoRI; Site 2: XhoI; Sized  
 fractionated cDNAs were directly ligated into pOT2.  
 Plasmid cDNA library."

## ORIGIN

Alignment Scores:  
 Pred. No.: 8.99e-33 Length: 508  
 Score: 480.00 Matches: 86  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 22.20% Indels: 0  
 DB: 14 Gaps: 0

US-09-813-329-6 (1-409) x CK135985 (1-508)

Qy 324 TyrTyrValTyrAlaGlnIleCysTyrAsnAsnSerHisSerGlnAsnGlyPheIleVal 343  
 |||||  
 Db 508 TATTACGTATACGCCAGATATGCTACAACTCGCACGACGAGGATTATCGTC 449  
 |||||  
 Qy 344 PheGlnGlyAspThrProPheLeuGlnCysLeuAsnThrValProThrAsnMetProHis 363  
 |||||  
 Db 448 TTTCAAGGAGACACTCCATTCTCGAGTGTGTTGAACACGGTCCCAACATGCCACAT 389  
 |||||  
 Qy 364 LysValHisThrCysHisThrSerGlyLeuIleHisLeuGluArgAsnGluArgIleHis 383  
 |||||  
 Db 388 AGGTGCACACTGCCACACGAGTGGTCTGATCCACTGGAACGAACGAGAGATTCAT 329  
 |||||  
 Qy 384 LeuLysAspIleHisAsnAspArgAsnAlaValLeuArgGluGlyAsnAsnArgSerTyr 403  
 |||||  
 Db 328 CTGAAGGACATTACACACGATCGCAATGCAGTTCCTCGCGGAGGGAACACCGAAGCTAC 269  
 |||||  
 Qy 404 PheGlyIlePheLysVal 409  
 |||||  
 Db 268 TTTGGCATCTTCAAGGTG 251

## RESULT 4

CK135525/c

LOCUS  
 DEFINITION

CK135525  
 CK135525.1 GI:38627461  
 EST.

ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE

ORGANISM  
 Drosophila melanogaster (fruit fly)

Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 Ephydroidea; Drosophilidae; Drosophila.

REFERENCE  
 1 (bases 1 to 532)  
 Stapleton,M., Brokstein,P., Hong,L., Tyler,D., Berman,B.,  
 Carlson,J., Champe,M., Chavez,C., Dorsett,V., Farfan,D., Frise,E.,  
 George,R., Gonzalez,M., Guarin,H., Harris,N., Li,P., Liao,G.,  
 Misra,S., Mungall,C.J., Nunoo,J., Pacleb,J., Paragas,V., Park,S.,  
 Phouanavong,S., Wan,K., Yu,C., Lewis,S.E., Celniker,S. and  
 Rubin,G.M.

TITLE  
 JOURNAL  
 COMMENT BDGP/HMI RH Drosophila EST Project  
 Unpublished (2001)  
 Other ESTs: RH51659.5prime  
 Contact: Stapleton, M.

BDGP  
 Lawrence Berkeley National Lab  
 One Cyclotron Rd, Berkeley, CA 94720, USA  
 Fax: 510 486 6798

Email: [http://www.fruitfly.org/EST\\_est@fruitfly.berkeley.edu](http://www.fruitfly.org/EST_est@fruitfly.berkeley.edu)  
 Based upon the presence of a XhoI site followed by a run of 14 or  
 more T residues at the beginning of the sequence, this clone was  
 polyadenylated. The resulting Poly-T sequence has been removed. hit  
 genomic AE003831: arm:2R [4876906,5137815]  
 estimated-cyto:46B13-46E4: 08/05/2002  
 Plate: RH.516 row: E column: 11  
 High quality sequence stop: 382.

FEATURES  
 source  
 1..532  
 /organism="Drosophila melanogaster"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:7227"  
 /clone="RH51659"  
 /sex="male and female"  
 /dev\_stage="Adult"  
 /lab\_host="DH5-alpha TonA"  
 /clone\_lib="RH Drosophila melanogaster normalized Head  
 pFlc-1"  
 /note="Organ: head; Vector: pFlc1; Site 1: XhoI; Site 2:  
 BamHI; Library was kindly generated by Piero Carninci at  
 the RIKEN. The library was normalized and excised using  
 Cre recombinase. Plasmid cDNA library."





```
/organism="Glossina morsitans morsitans"
/mol_type="mRNA"
/sub_species="morsitans"
/db_xref="taxon:37546"
/clone="Tse129f05.q1c"
/tissue_type="adult infected gut"
/clone_lib="Glossina morsitans morsitans adult infected
gut"
/notes="country: Zimbabwe; EST from adult gut infected with
T.brucei"

ORIGIN
Alignment Scores:
Pred. No.: 2,546-08 Length: 539
Score: 206.50 Matches: 54
Percent Similarity: 42.00% Conservative: 30
Best Local Similarity: 27.00% Mismatches: 49
Query Match: 9.55% Indels: 67
DB: 13 Gaps: 7

US-09-813-329-6 (1-409) x BX553466 (1-539)
QY 38 IleProLeuValLeuGlyPheIleGlyLeuGlyLeuValValAlaIleLeuAlaLeuThr 57
DB 10 TTGGCAATTGTTTGACTTTTGTA-----CTCTGGGTGCTCG 48
QY 58 IleTrp-----GlnThrThrArg 63
DB 49 ATCTGGCATTTAATCTTATACGNAATTAGACTCGCTACGAAATAATACGCACAGT 108
QY 64 ValSerHisLeuAspLysGluLeuLysSerLeuLysArgValValAspAsnLeuGln 83
DB 109 ATTAGAGATCTTCAAAGAGAGTCACTACACTCAACAGAGATCATTTGAAGTCTACAGAA 168
QY 84 ArgLeuGlyIleAsnTyrLeuAspGluPheAspGluPheGlnLysGluTyrGluAsnAla 103
DB 169 CGATTGGGTATACCATATCTGGATGATTTACATGATTTTGAAGAAGAGTATAAAATGTT 228
QY 104 Leu----- 104
DB 229 GTTGGAGCTTATCCCGCACCTTACGATTAAGTTTAAAGATAAAGATCTTTGACAAGAG 288
QY 105 -----IleAspTyrProLysLysVal-----AspGlyLeu 114
DB 289 AACGCAGAAAGTGAAGCGGAAGATTATGAGGAAGACATCGTTTACAGGAGGACGATCTC 348
QY 115 ThrAspGluLeuAspAspAspGlyAspGlyLeuAspSerIleAlaAspAspGluAsp 134
DB 349 AATGATAATGATAATGATAGTACGCGTCAT-----GATGACGATGAT 390
QY 135 AspAspValSerTyrSerValAspAspValGlyAlaAspTyrGluAspTyrThrAsp 154
DB 391 GATGATACCGATGACGGTAGGAGACGATGTTGAT---GACTATGAAGTTATGATGAT 447
QY 155 MetLeuAsnLysLeuAsnAlaHisThrGlyThrProThrSerGluThrThrAla 174
DB 448 TTAATGGAGAATTCGTAATTATCAGAAAT----- 477
QY 175 GluGlyGluGlyGluThrAspSerAlaSerSerAlaSerAsnAspAspAsnValPheAsp 194
DB 478 GAAGAGATAATAATAACGATGATAATGACTCGCGCGTCGTTAATGATGATAATTTATATGAC 537

RESULT 7
AI253821
LOCUS AI253821 526 bp mRNA linear EST 03-DEC-1999
DEFINITION AIMS-P.mon24 Giant tiger prawn eyestalk cDNA library Penaeus
monodon cDNA 5', mRNA sequence.
ACCESSION AI253821
VERSION AI253821.1 GI:3858749
KEYWORDS EST.
SOURCE Penaeus monodon (black tiger shrimp)
ORGANISM Penaeus monodon
Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
```

```
Eumalacostraca; Eucarida; Decapoda; Dendrobranchiata; Penaeoidea;
Penaeidae; Penaeus.
1 (bases 1 to 526)
Lehnert,S.A., Wilson,K.J., Byrne,K. and Moore,S.S.
Tissue-Specific Expressed Sequence Tags from the Black Tiger Shrimp
Penaeus monodon
Mar. Biotechnol. 1 (5), 465-476 (1999)
Contact: Wilson Kate J
Marine Biotechnology
Australian Institute of Marine Science
PMB 3, Townsville Mail Centre, Townsville QLD 4810, Australia
Tel: +61 7 4753 4444
Fax: +61 7 4772 5852
Email: kwilson@aims.gov.au
No significant database matches when searched using a filter to
remove low-complexity sequence, September 3 1998. When searched
with the filter turned off, shows homology to large proline-rich
protein bat2 from humans (p= 2.0e-07)
Insert Length: 1000 Std Error: 100.00
Seq primer: T3.

FEATURES
Location/Qualifiers
1..526
source
/organism="Penaeus monodon"
/mol_type="mRNA"
/db_xref="taxon:6687"
/sex="Female"
/tissue_type="Eyestalk"
/dev_stage="Mature broodstock"
/clone_lib="Giant tiger prawn eyestalk cDNA library"
/notes="vector: Lambda ZAP Express; Site 1: EcoRI; Site 2:
XhoI; mRNA isolated from total eyestalk tissue and used to
construct directional cDNA library in Lambda ZAP Express.
Clones excised in vivo to create phagemids in pBK-CMV for
sequencing."

ORIGIN
Alignment Scores:
Pred. No.: 3,55e-07 Length: 526
Score: 193.50 Matches: 59
Percent Similarity: 46.96% Conservative: 26
Best Local Similarity: 32.60% Mismatches: 70
Query Match: 8.95% Indels: 26
DB: 9 Gaps: 8

US-09-813-329-6 (1-409) x AI253821 (1-526)
QY 231 GlnGluLysSerSerAsnGluAlaThrSerLysGluSerProAlaProLeuHisHisArg 250
DB 45 CAAGAAGATCCCGAGGAGGAAGCAGAGTCA-----ACGCCGGCGCG-----CACCAA 92
QY 251 ArgArgMetHisSerArgHisArgHisLeuLeuValArgLysAlaArgSerGluAspSer 270
DB 93 CCCCGT-----GGTCACCGTCGCCCATCTCGTCGCCACGCCCGCCGACAGGAC--- 140
QY 271 ArgProAlaAlaHisPheHisLeuSerSerArg-ArgArgHisGlnGly---SerMetGl 289
DB 141 -----CGCCACCACCACGACGAGTAGCGAAGA 170
QY 289 YTyrHisGlyAspMetTyrIleGlyAsnAspAsnGluArgAsnSerTyrGlnGlyHisPhe 309
DB 171 CGTACACGGCGAGTGGTCTCTCGCGCTGGATGGACAAGTAGGTCTCAACCCCAAGTA 230
QY 309 eGlnThrArgAspGlyValLeuThrValThrAsnThrGlyLeuTyrTyrValTyrAlaGl 329
DB 231 CTCTCTCAGAGGGCGGTGTCACGGTCAAGGAATCTGGACTGTATTACCTCTATGCGCA 290
QY 329 nileCysTyrAsnAsnSerHisAspGlnAsnGlyPheIleValPheGlnGlyAspThrPr 349
DB 291 GGTGCTGTACAGCTGGAGCGTTCGCGAGTGGGTTCAGGTGGTCTCGACACGATCCC 350
QY 349 oPheLeuGlnCysLeuAsnThrValProThrAsnMetProHisLysValHisThrCysHi 369
DB 351 CATCATGAGTGC-----ACGTGGCTCCCGCTCAGCGCTCTCCT-----TCTTGCCA 398
```

```

QY 369 sThrSerGlyLeuIleHisLeuGluArgAsnGluArgIleHisLeuLysAspIleHisAs 369
Db 399 CACCGGGGGCCACCTACCTCTCCAGAACGTCGCGTACATCCGGGACCTGGACCA 458
QY 389 nAspArgAsnAlaValLeuArgGluGlyAsnAsnArgSerTyrPheGlyIlePheLysVa 409
Db 459 CCACATGACCGCGTGAAGATGAGGAGAAC-----AGCTTCTTCGGCCTCGTGAAGCT 512
QY 409 1 409
Db 513 G 513

RESULT 8
LOCUS CNS017WQ 1101 bp DNA linear GSS 26-JUL-1999
DEFINITION Drosophila melanogaster genome survey sequence T7 end of BAC
BACN37D20 of DrosBAC library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION AL108596
VERSION AL108596.1 GI:5628900
KEYWORDS GSS.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 1101)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT Determination of this BAC-end sequence was carried out as part of a
collaboration with the European Drosophila Genome Project (EDGP) -
http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC
library (Dros BAC) was made by Alain Billaud at CEPH (Centre
d'Etude du Polymorphisme Humain) with funding provided by a MRC
project grant. The DNA was prepared from embryos by Alain Bucheton
and Genevieve Payan. It has been constructed in the vector
pBelobAC11.

FEATURES             Location/Qualifiers
     source            1..1101
     organism="Drosophila melanogaster"
     mol_type="genomic DNA"
     db_xref="taxon:7227"
     clone="BACN37D20"
     clone_lib="DrosBAC"
     plasmid="pBelobAC11"
     note="end : T7"

ORIGIN
Alignment Scores:      8.22e-05      Length:      1101
Pred. No.:            172.00      Matches:      35
Score:                 67.21%      Conservative: 6
Best Local Similarity: 57.38%      Mismatches:  20
Query Match:          7.96%        Indels:       0
DB:                   29           Gaps:         0

US-09-813-329-6 (1-409) x CNS017WQ (1-1101)

QY 56 LeuThrIleTrpGlnThrArgValSerHisLeuAspLysGluLeuLysSerLeuLys 75
Db 256 TTGCAGAWMTGGCAGAAWNRGCGTGTTCGATCTGGACAAGAACTGAAACCTGAAV 315
QY 76 ArgValValAspAsnLeuGlnGlnArgLeuGlyIleAsnTyrLeuAspGluPheAspGlu 95
Db 316 CSARTCGTCRATAATCTCCAGCAACGTTTGGGCATAAACAATCTCGACGAATACGACGAT 375
QY 96 PheGlnLysGluTyrGluAsnAlaLeuIleAspTyrProLysLysValAspGlyLeuThr 115
Db 376 TTCAAAAARGAGGTCAGCAATTCGGTAAATCACTTTCCGAACATCAATATWWTACTCGGC 435

```

```

QY 116 Asp 116
Db 436 GAT 438

CNS07DXT 917 bp DNA linear GSS 08-JUL-2001
LOCUS T3 end of clone BD0AA016E09 of library BD0AA from strain CBS 94 of
DEFINITION Candida tropicalis, genomic survey sequence.
ACCESSION AL441063
VERSION AL441063.1 GI:12224469
KEYWORDS GSS.
SOURCE Candida tropicalis
ORGANISM Candida tropicalis
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetales; Mitosporic Saccharomycetales; Candida.
REFERENCE 1 (bases 1 to 917)
AUTHORS Souciet,J.L., Aigle,M., Artiguenave,F., Blandin,G.,
Bolotin-Fukuhara,M., Bon,E., Brottier,P., Casaregola,S.,
de-Montigny,J., Dujon,B., Durrens,P., Lepingle,A., Llorente,B.,
Malpertuy,A., Neuveglise,C., Ozier-Kalogeropoulos,O., Potier,S.,
Saurin,W., Tekala,F., Toffano-Nioche,C., Wesolowski-Louvel,M.,
Wincker,P. and Weissbach,J.
TITLE Genomic exploration of the hemiascomycetous yeasts: 1. A set of
Yeast species for molecular evolution studies
JOURNAL FEBS Lett. 487 (1), 3-12 (2000)
MEDLINE 20584711
PUBMED 11152876
REFERENCE 2 (bases 1 to 917)
AUTHORS Blandin,G., Ozier-Kalogeropoulos,O., Wincker,P., Artiguenave,F. and
Dujon,B.
TITLE Genomic exploration of the hemiascomycetous yeasts: 16. Candida
tropicalis
JOURNAL FEBS Lett. 487 (1), 91-94 (2000)
MEDLINE 20584726
PUBMED 11152891
REFERENCE 3 (bases 1 to 917)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (08-SEP-2000) Genoscope - Centre National de Sequencage,
2 rue Gaston Cremieux, CP 5706, 91057 EVRY cedex, FRANCE. (E-mail :
seqref@genoscope.cns.fr - Web :
This GSS is part of a random genomic sequencing program of thirteen
yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces
exiguus, Saccharomyces servazzii, Zygosaccharomyces rouxii,
Saccharomyces kluyveri, Kluyveromyces marxianus var. marxianus, Pichia
lactis var. lactis, Kluyveromyces fragilis, Debaryomyces hansenii var.
hansenii, Pichia sorbitophila,
Candida tropicalis and Yarrowia lipolytica. Genomic inserts of 3 to
5 kb were prepared and both extremities were sequenced. See
keywords for description of this sequence and for the sequence of
the other extremity of this insert.

FEATURES             Location/Qualifiers
     source            1..917
     organism="Candida tropicalis"
     mol_type="genomic DNA"
     strain="CBS 94"
     db_xref="taxon:5482"
     clone="BD0AA016E09"
     clone_lib="BD0AA"
     note="end : T3"

misc_feature
<10..>383
/note="similar to Saccharomyces cerevisiae ORF YKL088w [
similarly to C.tropicalis hal3 protein, to C-term. of
Sis2p and to hypothetical protein YOR054c ]
1 putative frameshift(s)"
/evidence=not_experimental

ORIGIN
Alignment Scores:      9.62e-05      Length:      917
Pred. No.:            170.00      Matches:      62
Score:

```

Percent Similarity: 42.79% Conservativity: 30  
 Best Local Similarity: 28.84% Mismatches: 73  
 Query Match: 7.86% Indels: 50  
 DB: 29 Gaps: 10

US-09-813-329-6 (1-409) x CNS07DXT (1-917)

```

Qy 40 LeuValLeuGlyPheIleGlyLeuGlyLeuValValAlaIleLeuAlaLeuThrIleTrp 59
Db 297 TTAGTTTGTGTGATATTGGAATGGA-----GGTATCGGAGATGG 338
Qy 60 GlnThrThrArgValSerHisLeuAspLysGluLeuLysSerLeu-----Lys 75
Db 339 GCTGATATT--GTCGATATCTTGAGAAGAAAAATCAATTCATTAAGCTGAACAAAAG 395
Qy 76 ArgValValAsnLeuGlnArgLeuGlyIleAsnTyrLeuAspGluPheAspGlu 95
Db 396 AGGTATTAGATGAA--GAAGAAGATTG-----AATGAAGAGGAGCAAGAGATGAG 446
Qy 96 PheGlnLysGluTyrGluAsnAlaLeuIleAspTyrProLysLysValAspGlyLeuThr 115
Db 447 GATGAAGAAGAGGGTGTGATGACGGC-----GAAGAAGATGGTGTGATCGG 488
Qy 116 AspGluGluAspAspAspGlyAspGlyLeuAspSerIleAlaAspAspGluAspAsp 135
Db 489 GACGATGATGATGATGACGACGATGATGATGAT-----GACGATGATGATGAT 539
Qy 136 AspValSerTyrSerSerValAspValGlyAlaAspTyrGluAspTyrThrAspMet 155
Db 540 GAC-----GACGACGAAGATGATGATGATGATGATGAAGAAGAAAT 581
Qy 156 LeuAsnLysLeuAsnAlaHisThrGlyThrThrProThrSerGluThrThrAlaGlu 175
Db 582 GATAACCAAAATACAAATGTGACATCGGT-----TCATCCAAAGGATTAGCTAAT 632
Qy 176 GlyGluGlyGluThrAspSerAlaSerSerAlaSerAsnAspAsnValPheAspAsp 195
Db 633 GGTAATCTGAAGCACCAGCAATCTAAAGATGACGAGAATGATGAATGATATTGAR--- 689
Qy 196 PheThrSerTyrAsnAlaHisLysLysGlnGluArgLysSerArgSerIleAlaAsp 215
Db 690 -----ATTACAGAT 698
Qy 216 ValArgAsnGluGluGlnAsnIleGlnGlyAsnHisThrGluLeuGlnGlySerSer 235
Db 699 CAAGGCAAAAGTGAGTCAGAAAGTCACAAACAGGATGAGAGATATCAAGATAGTTCATCA 758
Qy 236 AsnGluAlaThrSerLysGluSerProAlaProLeuHisHisArg 250
Db 759 ATTACTAAGAAACAACTGTTTCCCGGCACCACTCAACTCGTGT 803

```

# RESULT 10

BU373036  
 LOCUS BU373036 709 bp mRNA linear EST 28-NOV-2002  
 DEFINITION 603812126F1 CSEQCHN74 Gallus gallus cdna clone CHEST800c8 5', mRNA sequence.

ACCESSION BU373036

VERSION BU373036.1 GI:25881037

KEYWORDS EST.

SOURCE Gallus gallus (chicken)

ORGANISM Gallus gallus

REFERENCE 1 (bases 1 to 709)  
 Boardman, P.E., Sans-Ezquerro, J., Overton, I.M., Burt, D.W., Bosch, E.,  
 Fong, W.T., Tickle, C., Brown, W.R.A., Wilson, S.A. and Hubbard, S.J.  
 A Comprehensive Collection of Chicken cDNAs  
 Curr. Biol. 12 (22), 1965-1969 (2002)  
 22335534  
 12445392

AUTHORS Contact: Simon Hubbard

DEPARTMENT Department of Biomolecular Sciences

University of Manchester Institute of Science and Technology  
 (UMIST)  
 PO Box 88, Manchester, M60 1QD, UK  
 Tel: 01612008930  
 Fax: 01612360409  
 Email: Simon.Hubbard@umist.ac.uk.

## FEATURES

### source

1. 709  
 Location/Qualifiers  
 /organism="Gallus gallus"  
 /mol\_type="mRNA"  
 /strain="Compton Line 151"  
 /db\_xref="taxon:9031"  
 /clone="CHEST800c8"  
 /sex="Female"  
 /dev\_stage="adult"  
 /lab\_host="DH108"  
 /clone\_lib="CSEQCHN74"  
 /notes="Organ: kidney + adrenal; Vector: pBluescript II  
 KS(+); Site 1: EcoRI; Site 2: NotI; This normalized  
 library was constructed from 1 million independent clones.  
 cDNA synthesis was initiated using an oligo(dT) primer,  
 using methylated C in the first strand synthesis reaction.  
 Following this first strand reaction, double-stranded cDNA  
 was blunted, ligated to NotI adapters, digested with  
 EcoRI, size-selected, and cloned into the NotI and EcoRI  
 compatible sites of a custom modified MCS of the  
 pBluescript (KS+) vector. The library was normalized in 2  
 rounds using conditions adapted from Soares et al., PNAS  
 (1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6  
 (1996): 791, except that a significantly longer  
 reannealing hybridization was used."

## ORIGIN

Alignment Scores:  
 Pred. No.: 0.00334 Length: 709  
 Score: 151.00 Matches: 46  
 Percent Similarity: 49.34% Conservative: 29  
 Best Local Similarity: 30.26% Mismatches: 63  
 Query Match: 6.98% Indels: 14  
 DB: 13 Gaps: 7

US-09-813-329-6 (1-409) x BU373036 (1-709)

```

Qy 264 LysAlaArgSerGluAspSerArgProAlaAlaHisPheHisLeuSerSerArgArg 283
Db 146 AAGCGCGCTCCGAGACACCACCGACGCTGTG---GTCACCTTCAAGGC----- 193
Qy 284 HisGlnGlySer---MetGlyTyrHisGlyAspMetTyrIleGly-----AsnAspAsn 300
Db 194 ---CAAGGCTCAGCCATCCCAAGTGAAGATGATCTTTTCAGGTGGAGTCTCTCAATGACTGG 250
Qy 301 GlnArg-----AsnSerTyrGlnGlyHisPheGlnThrArgAspGlyValLeuThr 317
Db 251 TCTCGCATCATATGAACCAAGGTGTTTAAAGCTGCATGCCCGCAGTGGGAGTGAG 310
Qy 318 ValThrAsnThrGlyLeuTyrValTyrAlaGlnIleCysTyrAsnAsnSerHisAsp 337
Db 311 GTACTGGTGACGCGCACATATCTATCTATGATCAGGTATCTACTACATAAATCTTCACAGAT 370
Qy 338 GlnAsnGlyPheIleValPheGlnGlyAspThrProPheLeuGlnCysLeuAsnThrVal 357
Db 371 TTTGCCAGCTATGAGTGGTGGTGGATGAATAAACCCCTTTCTGCAATGCACTCGGAGTATT 430
Qy 358 ProThrAsnMetProHisLysValHisThrCysHisThrSerGlyLeuIleHisLeuGlu 377
Db 431 GAGACCGGCAAGACCAAC---TTCAATACCTGCTATACGCGCGGGGTCTGCCTCTCAAA 487
Qy 378 ArgAsnGluArgIleHisLeuLysAspIleHisAsnAspArgAsnAlaValLeuArgGlu 397
Db 488 GCCAGGCAAGAAATTGCTGTGAATAATGGTCCACGCTGACATCTCC-----ATCAACATG 541
Qy 398 GlyAsnAsnArgSerTyrPheGlyIlePheLeuVal 409
Db 398 GlyAsnAsnArgSerTyrPheGlyIlePheLeuVal 409

```

```

Db      542 AGCAAGCACACAACCTTTCTTTGGGCCATCGCCTA 577
RESULT 11
LOCUS   CF590046
DEFINITION AGENCOURT_15681475 NICHD_XGC_Swbin Silurana tropicalis cDNA clone
IMAGE:7020818 5', mRNA sequence.
ACCESSION CF590046
VERSION   CF590046.1 GI:36340024
KEYWORDS EST.
SOURCE   Silurana tropicalis (western clawed frog)
ORGANISM Silurana tropicalis
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
          Xenopodinae; Silurana.
REFERENCE 1 (bases 1 to 965)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE   National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Daniela S. Gerhard, Ph.D.
          National Cancer Institute / NIH
          Bldg. 31 Rm10A07 Bethesda, MD 20892
          Email: cgapbs-remail.nih.gov
          Tissue Procurement: Rob Granger, University of Virginia
          cDNA Library Preparation: Open Biosystems
          DNA Sequencing by: Agencourt Bioscience Corporation
          Clone distribution: MGC clone distribution information can be
          found through the I.M.A.G.E. Consortium/LLNL at:
          http://image.llnl.gov
          Plate: L1AM14744 row: g column: 24
          High quality sequence start: 23
          High quality sequence stop: 511.
FEATURES             source
     source
     1..965
     /organism="Silurana tropicalis"
     /mol_type="mRNA"
     /db_xref="taxon:8364"
     /clone="IMAGE:7020818"
     /tissue_type="whole body"
     /notes="Vector: pExpress-1; Site 1: EcoRV; Site 2: NotI;
     Bulk tissue was collected from a whole 10 month old male
     from the F6 strain. 1st strand cDNA was primed with a Not
     I - oligo(dT) primer, double-stranded cDNA was cloned into
     the Not I and EcoRV sites of pExpress-1. Library was
     size-selected for >1.5 kb fragments for an average insert
     size of 1.92 kb. Library was normalized to Cot5 with a
     180-fold reduction of actin. A non-normalized version of
     this library is also available (NICHD_XGC_Swbi). Library
     was constructed by Open Biosystems (Huntsville, AL)."
ORIGIN
Alignment Scores:
Pred. No.:      0.0063      Length:      965
Score:          150.00      Matches:    47
Percent Similarity: 46.24%      Conservat: 33
Best Local Similarity: 27.17%      Mismatches: 60
Query Match:     6.94%      Indels:    33
DB:              14      Gaps:      7

US-09-813-329-6 (1-409) x CF590046 (1-965)

QY      232 GluLysSerSerAsnGluAlaThrSerLysGluSerProAlaProLeuHisHisArgArg 251
Db      148 GAGAAAGAGTCACAGAGAACCGAGGCCAGAGGTGAGAGACCAACCAATACAAAGAAGA 207
QY      252 ArgMetHisSerArgHisArgHisLeuValArgLysAlaArgSerGluArgSerArg 271
Db      208 AGGCCTGGCAAT-----CTGTCTCAGACAGAAAGAAGA 240
QY      272 ProAlaAlaHisPheHisLeuSerSer-ArgArgArgHisGlnGlySerMetGlyTyr-- 290

```

```

Db      241 CCT-----CACGGAGCCAGAGCCAGACATCACGGAACCGTTTCAATATGT 288
QY      291 -His-----GlyAspMetTyrIleGlyAsnAspAsn----- 300
Db      289 TCACCTGGTGCTGATCATTTCTTTACCAATAGAGCCCTGGATGGACCTCTCGTCTCCTG 348
QY      301 -GluArgAsnSerTyrGlnGlyHis---PheGlnThrArgAspGlyValLeuThrValTh 319
Db      349 GAAAGTGTCTATGAAGGAAGGCGAGCATTTCAAGGTTCAAGGTGTTAAATGTGACAGTGAG 408
QY      319 rAsnThrGlyLeuTyrValTyrAlaGlnIleCysTyrAsnAsnSerHisAspGlnAs 339
Db      409 GAATTCGGGATATATTCAGTATACAGCAGGTCTGTACACAGATTCAGGTTTCAGAT 468
QY      339 nglyPheIleValPheGln-----GlyAspThr-ProPheLeuG1 352
Db      469 GGGACACTTGATAACACGTAAGCTAGAGGGCTCCAGGAGGAGAGAAGTGCTGTACG 528
QY      352 nCysLeuAsnThrValProThrAsnMetProHisLysValHisThrCysHisThrSerG1 372
Db      529 CTGTGTGCAGAGCATGCCGCCCAATGAGAGCATGGCGTATACACCTGTTACAGTCCCG 588
QY      372 YLeulleHisLeuGluArgAsnGluArgIleHisLeu 384
Db      589 AGTTTCCAACTGCAAAAAGGGGATTCATTAGCCTC 625

RESULT 12
LOCUS   CD851068
DEFINITION DH0AL52H012M1 HaDevR5 Helianthus annuus cDNA clone HaDevR55H01,
mRNA sequence.
ACCESSION CD851068
VERSION   CD851068.1 GI:32534884
KEYWORDS EST.
SOURCE   Helianthus annuus (common sunflower)
ORGANISM Helianthus annuus
          Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
          Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
          asterids; campanulids; Asterales; Asteraceae; Asteroideae;
          Heliantheae; Helianthus.
REFERENCE 1 (bases 1 to 746)
AUTHORS Genoplante.
TITLE   Genoplante, a major partnership french program in plant genomics
JOURNAL Unpublished (2003)
COMMENT Contact: Genoplante
          Genoplante
          93, rue Henri Rochefort 91025 EVRY CEDEX France
          Tel: 33 1 69 47 54 00
          Fax: 33 1 69 47 54 10
          This sequence has been generated in the framework of the french
          plant genomics programme 'Genoplante' (http://www.genoplante.com
          and http://genoplante-info.infobiogen.fr).
FEATURES             source
     source
     1..746
     /organism="Helianthus annuus"
     /mol_type="mRNA"
     /cultivar="pssc8"
     /db_xref="taxon:4232"
     /clone="HaDevR55H01"
     /tissue_type="4 days after self-pollination embryo"
     /clone_lib="HaDevR5"
ORIGIN
Alignment Scores:
Pred. No.:      0.0123      Length:      746
Score:          145.00      Matches:    47
Percent Similarity: 43.53%      Conservat: 27
Best Local Similarity: 27.65%      Mismatches: 48
Query Match:     6.71%      Indels:    48
DB:              14      Gaps:      7

US-09-813-329-6 (1-409) x CD851068 (1-746)

```



```

MEDLINE      22879925
COMMENT      Contact: Zeguang Han
             Chinese National Human Genome Center at Shanghai
             351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
             201203, P. R. China
             Tel: 86-21-50801919(ex.45)
             Fax: 86-21-50801922
             Email: hanzg@chgc.sh.cn.

FEATURES     source
             1..724
             /organism="Schistosoma japonicum"
             /mol_type="mRNA"
             /db_xref="taxon:6182"
             /sex="female"
             /tissue_type="whole body"
             /dev_stage="adult"
             /lab_host="rabbits"
             /clone_lib="SJF"

ORIGIN
Alignment Scores:
Pred. No.:      0.0178      Length:      724
Score:          143.00      Matches:    45
Percent Similarity: 43.26%  Conservative: 32
Best Local Similarity: 25.28% Mismatches:   79
Query Match:     6.61%     Indels:     22
DB:              13        Gaps:       5

US-09-813-329-6 (1-409) x BU798281 (1-724)
QY  91 AspGluPheAspGluPhe-----GlnLysGluTyrGluAsnAlaLeuLeuAspTyr 107
   |||||
Db  124 GATGAATTGATGAGTTCGGTCGTAAGAAAGAAAGTATCGCGCAAGTCAAGTGAACAA 193
   |||||

QY  108 ProLysLysValAspGlyLeu-----
   |||||
Db  184 CTTAGTAACCATGACTCGTCGGATTCTCGTAACAATAATGCTACACCAAGACATGATCAT 243
   |||||

QY  115 ThrAspGluGluAspAspGlyAspGlyLeuAspSerIleAlaAspAspGluAsp 134
   |||||
Db  244 AATGATGAATAATGCAATAACAAGGATGATGATGATGATGATGATGATGATGATGATGATG 303
   |||||

QY  135 AspAspValSerTyrSerValAspValGlyAlaAspTyrGluAspTyrThrAsp 154
   |||||
Db  304 GATGATGCTGATCTCTTAAGTATGATATTGGGGGGCAGAAATGACGATTTAGGTGAC 363
   |||||

QY  155 MetLeuAsnLysLeuAsnAsnAlaHisThrGlyThrThrProThrSerGluThrAla 174
   |||||
Db  364 -----AAGTTATCAAAATGCACAA-----AATCTCGTATCTTCTCAAGTATCTTACC 408
   |||||

QY  175 GluGlyGluGlyGluThrAspSerAlaSerAlaSerAlaSerAspAspAspAspValPheAsp 194
   |||||
Db  409 AATAAAAAATCTCGAAGTCGTAGTCCATGCTCTTCCTCGGACTCAGACTCAGATTTCTTCC 468
   |||||

QY  195 AspPheThrSerTyrAsnAlaHisLysLysLysGlnGluArgLysSerArgSerIleAla 214
   |||||
Db  469 TCATCATCCAGCCAGACTTCACATTCCTCAACATCAGATTCATCATCAAGTTCTCGTAGT 528
   |||||

QY  215 AspValArgAsnGluGluGlnAsnIleGlnGlyAsnHisThrGluLeuGlnGluLysSer 234
   |||||
Db  529 GAGTCATCTTCCAAACCCGAAAAAT---CAATCCAAACCGGCAAGAGACCTCATCTCAAAAAAT 585
   |||||

QY  235 SerAsnGluAlaThrSerLysGluSerProAlaProLeuHisHisArgArgGag 252
   |||||
Db  586 GGCATATCGGGCAACGATAGTACAGTTCACCTACCGTAATACCAAGCGAAAA 639
   |||||

RESULT 15
BU803616/c
LOCUS      BU803616
DEFINITION SJFAWE05 SJF Schistosoma japonicum cDNA, mRNA sequence.
ACCESSION BU803616
VERSION   BU803616.1 GI:28360623
KEYWORDS  EST.

SOURCE      Schistosoma japonicum
ORGANISM    Schistosoma japonicum
REFERENCE   1 (bases 1 to 732)
AUTHORS     Hu, W., Yan, Q., Shen, D.K., Liu, F., Zhu, Z.D., Song, H.D., Xu, X.R.,
             Wang, Z.J., Rong, Y.P., Zeng, L.C., Wu, J., Zhang, X., Wang, J.J.,
             Xu, X.N., Wang, S.Y., Fu, G., Zhang, X.L., Wang, Z.Q., Brindley, P.J.,
             McManus, D.P., Xue, C.L., Feng, Z., Chen, Z. and Han, Z.G.
             Evolutionary and biomedical implications of a Schistosoma japonicum
             complementary DNA resource
             Nat. Genet. 35 (2), 139-147 (2003)
             22879925
             Contact: Zeguang Han
             Chinese National Human Genome Center at Shanghai
             351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
             201203, P. R. China
             Tel: 86-21-50801919(ex.45)
             Fax: 86-21-50801922
             Email: hanzg@chgc.sh.cn.
             Location/Qualifiers
             1..732
             /organism="Schistosoma japonicum"
             /mol_type="mRNA"
             /db_xref="taxon:6182"
             /sex="female"
             /tissue_type="whole body"
             /dev_stage="adult"
             /lab_host="rabbits"
             /clone_lib="SJF"

ORIGIN
Alignment Scores:
Pred. No.:      0.0181      Length:      732
Score:          143.00      Matches:    45
Percent Similarity: 43.26%  Conservative: 32
Best Local Similarity: 25.28% Mismatches:   79
Query Match:     6.61%     Indels:     22
DB:              13        Gaps:       5

US-09-813-329-6 (1-409) x BU803616 (1-732)
QY  91 AspGluPheAspGluPhe-----GlnLysGluTyrGluAsnAlaLeuLeuAspTyr 107
   |||||
Db  593 GATGAATTGATGAGTTCGGTCGTAAGAAAGAAAGTATCGCGCAAGTCAAGTGAACAA 534
   |||||

QY  108 ProLysLysValAspGlyLeu-----
   |||||
Db  533 CTTAGTAACCATGACTCGTCGGATTCTGTAACAATAATGCTACACCAAGACATGATCAT 474
   |||||

QY  115 ThrAspGluGluAspAspAspGlyAspGlyLeuAspSerIleAlaAspAspGluAsp 134
   |||||
Db  473 AATGATGAATAATGCTCAATAACAAGGATGATGATGATGATGATGATGATGATGATGATGATG 414
   |||||

QY  135 AspAspValSerTyrSerValAspValGlyAlaAspTyrGluAspTyrThrAsp 154
   |||||
Db  413 GATGATGCTGATCTCTTAAGTATGATATTGGGGGGCAGAAATGACGATTTAGGTGAC 354
   |||||

QY  155 MetLeuAsnLysLeuAsnAsnAlaHisThrGlyThrThrProThrSerGluThrAla 174
   |||||
Db  353 -----AAGTTATCAAAATGCACAA-----AATCTCGTATCTTCTCAAGTATCTTACC 309
   |||||

QY  175 GluGlyGluGlyGluThrAspSerAlaSerAlaSerAlaSerAspAspAspAspValPheAsp 194
   |||||
Db  308 AATAAAAAATCTCGAAGTCGTAGTCCATGCTCTTCCTCGGACTCAGACTCAGATTTCTTCC 249
   |||||

QY  195 AspPheThrSerTyrAsnAlaHisLysLysLysGlnGluArgLysSerArgSerIleAla 214
   |||||
Db  248 TCATCATCCAGCCAGACTTCACATTCCTCAACATCAGATTCATCATCAAGTTCTCGTAGT 189
   |||||

QY  215 AspValArgAsnGluGluGlnAsnIleGlnGlyAsnHisThrGluLeuGlnGluLysSer 234
   |||||
Db  188 GAGTCATCTTCCAAACCCGAAAAAT---CAATCCAAACCGGCAAGAGACCTCATCTCAAAAAAT 132
   |||||

```

Qy 235 SerAsnGluAlaThrSerLysGluSerProAlaProLeuHisHisArgArg 252  
Db 131 GGCAATACGGGCAACGATAGTGACAGCTCACCTACCGTAAATCACAGCGAAA 78

Search completed: September 15, 2004, 11:25:28  
Job time : 4198 secs

